OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: March 2, 2005, 06:00:31 ; Search time 3385 Seconds
                                         (without alignments)
                                         9175.907 Million cell updates/sec

Title: US-10-720-018-1
Perfect score: 816
Sequence: 1 atgactgactacattccaac.........atggtggttcacttcttgg 816

Scoring table: IDENTITY NUC
               Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 1903214700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   ;EST:*
               1: gb_est1:*
               2: gb_est2:*
               3: gb_htc:*
               4: gb_est3:*
               5: gb_est4:*
               6: gb_est5:*
               7: gb_est6:*
               8: gb_gss1:*
               9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Match	Query Match Length	DB	ID	Description
1		276.6	33.9	899	9 6	CNS077V0	AL433186 T7 end of
2		129.2	15.8	969	9 6	CD457882	CD457882 Fg04d.02e
3		118.4	14.5	682	4	BI750180	BI750180 Fg02_10b0
4	c	113.4	13.9	713	7	CF677471	CF677471 CCAI990TF
5	c	113.4	13.9	836	7	CF701403	CF701403 CCACT16TF
6	c	113.4	13.9	836	7	CF715948	CF715948 CCAB241TF
7	c	112.8	13.8	777	7	CF817896	CF817896 EST695278
8	c	112.8	13.8	787	7	CF824542	CF824542 EST701924
9	c	111.2	13.6	738	7	CF812382	CF812382 EST689764
10	c	106.4	13.0	676	7	CF824543	CF824543 EST701925
11	c	102.8	12.6	718	7	AJ638396	AJ638396 EST830966
12	c	100.8	12.4	615	7	CO136295	CO136295 EST835926
13	c	97.2	11.9	447	7	CO141255	CO141255 EST835926
14	c	93.2	11.4	511	7	AU249728	AU249728 AU249728
15	c	90.8	11.1	892	7	CN811863	CN811863 Fg09_08p2
16	c	82.8	10.1	753	5	CN812466	CN812466 Fg11_01e0
17	c	77.4	9.5	896	7	BQ855471	BQ855471 QGB26L21.
18	c	77.2	9.5	574	6	CF820440	CF820440 EST677822
19	c	76.6	9.4	833	7	CD456321	CD456321 Fg03_03g0
20	c	76	9.3	839	7	CO012227	CO012227 EST800562
21	c	76	9.3	922	7	CO012226	CO012226 EST800561
22	c	76	9.3	949	7	CO009933	CO009933 EST798268
23	c	76	9.3	963	7	CO004456	CO004456 EST792791
24	c	76	9.3		7	CO004455	CO004455 EST792790
25		75.8	9.3	692	5	BQ992634	QGF26C18.
26		74.8	9.1	809	6	CB905184	tric041x1
27		74.4	9.1	809	6	CF876511	tric041x1
28		73.8	9.0	479	7	CF191935	12o21j2.f
29	c	73	9.0	816	2	BE642336	Cri2_5_M0
30	c	72.6	8.9	688	5	BP098973	BP098973
31	c	72.6	8.9	735	7	CO140332	EST835003
32	c	71.4	8.7	912	7	CO009934	EST798269
33	c	70.8	8.7	690	7	CK569564	HO13K23w
34	c	70.8	8.7	1048	9	AL433069	T3 end of
35	c	70.2	8.6	712	6	CD458386	Fg08_12k2
36	c	70	8.6	482	7	CF190316	k8a22j2.f
37	c	68	8.3	734	7	CK447731	N12G9.SP6
38	c	67.4	8.3	659	1	AJ638576	AJ638576
39	c	67.4	8.3	971	9	AL439697	T3 end of
40	c	66.8	8.2	800	7	CO006699	EST799034
41	c	65.4	8.0	786	7	BQ855467	QGB26L16.
42	c	65.2	8.0	868	7	CO006700	EST795035
43	c	64.6	7.9	700	7	CK567406	HO06K05w
44	c	63.8	7.8	621	5	BU014570	QGJ7L07.Y
45	c	63.6	7.8	703	5	BQ861127	QC17J18.

ALIGNMENTS

RESULT 1
CNS077V0

LOCUS   CNS077V0   899 bp   DNA   linear   GSS 08-JUL-2001
DEFINITION T7 end of clone BB0AA014G11 of library BB0AA from strain CBS 4732
            of Pichia angusta, genomic survey sequence.
ACCESSION   AL433186
VERSION   AL433186.1 GI:12216600
KEYWORDS   GSS.
SOURCE   Pichia angusta
  ORGANISM Pichia angusta
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE   1 (bases 1 to 899)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukihara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuvegliae,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE   Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2 (bases 1 to 899)
  AUTHORS   Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F.
            and Dujon,B.
  TITLE   Genomic exploration of the hemiascomycetous yeasts: 13. Pichia
            angusta
  JOURNAL   FEBS Lett. 487 (1), 76-81 (2000)
  MEDLINE   20584723
   PUBMED   11152888
REFERENCE   3 (bases 1 to 899)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT   This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of

RESULT 1

```
FEATURES             Location/Qualifiers
     source          1..899
                     /organism="pichia angusta"
                     /mol_type="genomic DNA"
                     /strain="CBS 4732"
                     /db_xref="taxon:4905"
                     /clone="BB0AA014G11"
                     /clone_lib="BB0AA"
                     /note="the other extremity of this insert."
     misc_feature    <18..>633
                     /note="similar to P50167 [ D-arabinitol 2-dehydrogenase
                     ARDH | Pichia stipitis ]
                     1 putative frameshift(s)"
                     /evidence=not experimental
     misc_feature    complement(<729..>812)
                     /note="similar to Saccharomyces cerevisiae ORF YKL126w [
                     YPK1 ; ser/thr-specific protein kinase ]"
                     /evidence=not_experimental
     misc_feature    complement(<732..>812)
                     /note="similar to Saccharomyces cerevisiae ORF YMR104c [
                     YPK2 ; ser/thr protein kinase ]"
                     /evidence=not_experimental
ORIGIN

Query Match      33.9%;  Score 276.6;  DB 9;  Length 899;
Best Local Similarity  66.0%;  Pred. No. 3.6e-68;
Matches  416;  Conservative  0;  Mismatches  210;  Indels  4;  Gaps  1;

Qy    191 TGAAGTTGAAAGAAGTTCCAAAGATGGGGTTCATTAGCCTGTGATAATTTCTGATTCTGATA 250
          ||||  ||||||||  | ||| |||||||| |||||||||  || | |||  ||| ||| ||
Db      4 TCAAGGAGGAGATCTCGTGCTCAAAAGATGCACTCGTACGTTTGCGACATTAGCGAGCGGCTGAGC 63

Qy    251 CCGTTCACAAGGTGTTTGCTCAAGTTGCTAAGGTAAGTTGCCATTGCCACTTGG 310
          |||||| ||| ||| ||||| |||||| ||||||| ||| ||||||| ||||||
Db     64 ACGTGAACGAGGGTCTTGCTCAAATATACGGCCAGCTTTGGAAAGTACCCGCTGCATCTCG 123

Qy    311 TTAACACAGCTGGTTACTGTGAAAACTTCCCATGTGAAGATTACCCAGCCAAGAACCGCTG 370
          ||| || || |||| || || ||||||||||| ||| ||| |||| ||| ||| ||| |
Db    124 TCAACACTGCTGGTTATTGCGAGAATTTCCCACGCCCCCCAAGAACGCCG 183

Qy    371 AGAAGATGGTGAAGGTTAACTTGTTGGGGTTCTTTGTATGTTTCTCAAGCCTTTGCTAAGC 430
          |||| | |||| || ||||| ||| ||||||| || || ||| |||| | |||| |||||
Db    184 AAAAGCTTCTCAAGGTCAACCTTTTGGGCGCCTTGTTGTTACGTTGCACAAGCATTCGCCAAAC 243

Qy    431 CATTGATCAAAAGAGAAGAGTATCAAGGGGTGCTTCTGTTGTTGGTTACGTTCTATGTCTGGTG 490
          |||| | || |||||||||| |||| |||||| |||| || ||||| ||| ||| ||| |||
Db    244 CACTTATCGTATACAACATCAAGGAGGAGGCTCCATCGGATCCATGTCCATGTCCGGAG 303

Qy    491 CCATTGTCAACGATCCTCAAAACCAAGTTGTCTACAACACATTGTCAACTGTGATGATTCTTATCC 550
          || ||| || |||||| ||| || |||| ||||||| || |||| ||||| || ||| |||||
Db    304 AAATCGTCAAACGACCCACAAACCGAGTCGCCTACACAAACGCGGTGTGTCTTCTTATC 363

Qy    551 ATTTGGCTAAGACTTTGGCTTGTGTGAATGGGGTAAGTACAACATCAGCAGTTAATTCTTTAA 610
          ||||||| |||||||| ||||| || ||||| |||| ||| |||||||| ||| |||| ||
Db    364 ACATGGTCAAGTCTTCTCCGCCGCTGAAATGGGGCCAAATACAGCATCAGAGTGAAATACCTTGT 423

Qy    611 ACCCAGGTTACATCTACGGTCCTTGACCAAGATGTTATCAATGGGAGGAAGAAATGT 670
          ||||||||||| ||||| |||||||||||| || |||||  |||| || || |||
Db    424 CGCCAGGATCATCCTTCACACCGCTCACAAAAGAAGAGCGTTATTAACGGATAATAACGGAAAAGT 483

Qy    671 ACAACAGAGAAGGATCTCTGGTATCCCACAACAAAAGGATACCACAATTG 730
          ||| |||| ||||| |||| ||| | ||| || ||||| ||||| |||
Db    484 ACCAGAGAGATGGCTTTGCCGGTCAAGTCTTTCGGAAGCTAAGGAAGATTCACACG 543

Qy    731 GTGCGTGTTTGTACTTGCTTTCTGAATCTGCTGCTTCATACACTACTTGGTCTGCCAGCTTA 789
          ||||||  ||||| ||||||||||| || |||| || ||| ||| ||| || ||| ||||
Db    544 GTACGTNTCTACTTGCTCTCCAACTCCGCTTCTTCAGCATTTACTACCACGGGAGCAAATCAT 603

Qy    790 ---CTGGTTGATGGTTGGTTTTCACTTCTTGG 816
             || |||||| |||||| |||| |||| ||
Db    604 TATTTGTCTGACGGTGGTTTTCACTTCTTCTGG 633
```

RESULT 2

```
CD457882
LOCUS       CD457882                969 bp    mRNA    linear   EST 14-JUN-2004
DEFINITION  Fg04d_02e10_A Fg04_AAFC_ECORC_Fusarium_graminearum mycelium grown_on_wheat_heads
            Gibberella zeae cDNA clone Fg04d_02e10, mRNA sequence.
ACCESSION   CD457882
VERSION     CD457882.1  GI:31372622
KEYWORDS    EST.
SOURCE      Gibberella zeae
  ORGANISM  Gibberella zeae
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE   1  (bases 1 to 969)
  AUTHORS   Ouellet,T., Kou],A., Dan,H., Harris,L.J., Chapados,J., Couroux,P.,
            De Moors,A., Hattori,J., Lacroix,C., Masotti,M., Robert,L.S.,
            Singh,J.A., Sprott,D. and Tinker,N.A.
  TITLE     Fusarium graminearum mycelium grown on wheat heads under high
            humidity conditions
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ouellet, Therese
            Agriculture and Agri-food Canada
            Eastern Cereal and Oilseed Research Centre
            Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
            CANADA
            Tel: (613) 759-1658
            Fax: (613) 759-1701
            Email: ouellett@agr.gc.ca.
FEATURES             Location/Qualifiers
     source          1..969
                     /organism="Gibberella zeae"
                     /mol_type="mRNA"
                     /strain="DAOM 180378"
                     /db_xref="taxon:5518"
                     /clone="Fg04d_02e10"
                     /tissue_type="Mycelium"
                     /dev_stage="Asexual"
                     /lab_host="E. coli"
                     /clone_lib="Fg04_AAFC_ECORC_Fusarium_graminearum mycelium
                     grown_on_wheat_heads"
                     /note="Vector: Bluescript SK+; Site_1: EcoRI; Site_2:
                     XhoI; Fusarium grown on wheat (cv. Roblin) under high
                     humidity. cDNA made using Stratagene kit."
ORIGIN

Query Match      15.8%;  Score 129.2;  DB 6;  Length 969;
Best Local Similarity  57.1%;  Pred. No. 1.1e-25;
Matches  289;  Conservative  3;  Mismatches  206;  Indels  8;  Gaps  3;

Qy    305 ACTTGGTTAACACAGCTGGTTACTGTGAAAACTTCCCATGTGAAGATTACCCAGCCAAGA 364
          |||||| ||| ||| ||||| ||||| || ||||||| | |||||| |||| ||||||||
Db    254 ACTTGGTCACCTCAGCTGCTTCACGGAGAACTTCGAGCCGTTAACTACCCCATCGACC 313

Qy    365 ACGGTGAAGAGATGGGTGAAGGTTAACTTGTTGGGTTCTTTGTATGTTTCTCAAGCCTTTG 424
          ||||| |||| ||||| ||| ||||||  ||| ||||| |||| | ||| ||||||||||
Db    314 GTCTCGGTAAGCTTCGGGCTGTTAACGTTGACACGGTACATATCTCTTTGCAACATCAGTCG 373

Qy    425 CTAAGCCANTGAT-CAAAGAAGGTATCAAGGGGTCCTTCGTGTTTGTTGATTGGTTCTATG 483
          ||||||| | |||  ||||| || ||||||| ||| | ||||| || || |||| ||||
Db    374 CCAGGCACTTGATGCAAAGAAAAGCTCCTGGTAGCATC---GTCATGATTGGTTAGCATG 429

Qy    484 TCTGGTGCCATTGTCAACGATCCTCAAAACCAAGTTGTCTACAACACATTGTCAACAGTG 543
          ||| || || ||||| ||||||| || || ||||| ||||| | ||||||| ||| |||
Db    430 TCCGGATCCATTGTCAACGATCCTCASCCTCAGCGTCCTATAATGCCCAAAGCCGGT 489

Qy    544 GTTATCCCATTTGGCTAAGACTTTGGCTTGTGTGAATGGGGTAAGTACAACATCAGAGTTAAT 603
          ||| || ||||||| ||||| ||||| ||| || |||||| || ||| |||| ||| |||
Db    490 GTGCGCCCATCTCGCTCTTCCTTTGGCCGTCGAATGGGGCTCAGGCAAACATCCGGAGTCAAC 549

Qy    604 TCTTTAAACCCAGGTTACATCTACCAGGGTCTTGACCAAGAAGAATGTTATCAATGGTAACGAA 663
          ||||||||| ||||||  |||| |||||||| |||||||| || || ||||| ||||| |||
Db    550 TGCATCTCTCCCGGTTACATGTTGACTCCACTCAGAGAGAATTCTTGACGACAAACCCG 609
```

RESULT 3

```
LOCUS       BI750180               682 bp    mRNA    linear    EST 14-JUN-2004
DEFINITION  Fg02_10b03_R Fg02_AAFC_ECORC Fusarium_graminearum mycelium
            Gibberella_zeae cDNA clone Fg02_10b03, mRNA sequence.
ACCESSION   BI750180
VERSION     BI750180.1  GI:15771982
KEYWORDS    EST.
SOURCE      Gibberella zeae
  ORGANISM  Gibberella zeae
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   Harris,L.J., Glassco,T., Rocheleau,H., Allard,S., Chapados,J.,
            Couroux,P., De Moors,A., Hattori,J.I., Ouellet,T., Robert,L.S.,
            Singh,J.A., Sprott,D. and Tinker,N.A.
  TITLE     Expressed Sequence Tags from Fusarium graminearum mycelium
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Harris, Linda J.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-Food Canada
            Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
            CANADA
            Tel: (613) 759-1314
            Fax: (613) 759-6566
            Email: harrislj@agr.gc.ca.
FEATURES            Location/Qualifiers
     source          1..682
                     /organism="Gibberella zeae"
                     /mol_type="mRNA"
                     /strain="DAOM 180378"
                     /db_xref="taxon:5518"
                     /clone="Fg02_10b03"
                     /tissue_type="Mycelial tissue"
                     /dev_stage="Asexual"
                     /lab_host="E. coli (Sure cells)"
                     /clone_lib="Fg02_AAFC_ECORC Fusarium_graminearum mycelium"
                     /note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
                     Site_2: XhoI; Mycelial tissue was collected from V8 agar
                     plates after a growth period of 6-7 days at 25 C with 14
                     hrs (FL/UV) day ligth exposure. Mycelia was ground in
                     liquid nitrogen prior it's storage at -80 C until RNA
                     extraction. Directional cloning with 5' end of cDNA cloned
                     into EcoRI site of pBluescript and 3' end of cDNA cloned
                     into XhoI site of pBluescript (Stratagene,La Jolla, CA)."
ORIGIN
```

Qy   664 GAATTGTACAACACGATGGATCTCTGGGTATCCCAACAAAAGAATGTCCGAACCAAAGGAA 723
           ||||| ||||||||||||||||||||||||||||| ||| ||||||||| ||||||||||
Db   610 GACTCAAGGCCAAGTGGACTTCCCTTATCCCCAGGGCAAAAATGGGACAACCACACAGGAC 669

Qy   724 TACATTGGTGTCGTGTTTTGTTGTACTTGCTTTCTGAATCTGCTCCATTACACTACTGGTT 783
           |||||||||||||||||||||| |||||| |||||||||| |||| ||||||||||||
Db   670 CTCATGGGGGCCCCGTGGGCATYCCYCCTATCAGA---TGCTTCTTCGTATGTGACTGGTGCC 726

Qy   784 AGCTTACTGGTTGATGGTGGTTTCAC 809
                   ||||| |||||||||||||
Db   727 CACATCCGASTCGAGGCSGCTACAC 752

RESULT 4

```
LOCUS       CF677471               713 bp    mRNA    linear    EST 16-AUG-2004
DEFINITION  CCAI990TF C.neoformans strain JBC21 Cryptococcus neoformans var.
            neoformans cDNA clone CCA1990, mRNA sequence.
ACCESSION   CF677471
VERSION     CF677471.1  GI:41531630
KEYWORDS    EST.
SOURCE      Cryptococcus neoformans var. neoformans
  ORGANISM  Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
            var. neoformans)
            Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
            Heterobasidiomycetidae; Tremellales; Tremellaceae;
            Filobasidiella.
REFERENCE   1  (bases 1 to 713)
  AUTHORS   Loftus,B.
  TITLE     End sequencing of clones from a Full length enriched, normalized
            JBC21 cDNA library
  JOURNAL   Unpublished (2003)
  COMMENT   Other ESTs: CCA1990TR
            Contact: Brendan Loftus
            TIGR
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-3543
            Fax: 301-838-0208
            Email: crypt@tigr.org
            Seq primer: TF.
FEATURES            Location/Qualifiers
     source          1..713
                     /organism="Cryptococcus neoformans var. neoformans"
                     /mol_type="mRNA"
                     /strain="JBC21"
                     /db_xref="taxon:40410"
                     /clone="CCA1990"
                     /clone_lib="C.neoformans strain JBC21"
```

Qy   134 ACCAAGAAAAGACTTGCTGCCAAAAAAGAAGCCGAATACCACACAAATA---CGCTACTGNAGAAT 190
           ||||||||| |||||||||||||| |||||||||||||||||||| || ||||| ||||||
Db   140 ATAAGGAAGAGCAGAGAAGCAGACTAGCTTGATCATTGAAGAGTTCAAAAGGAGAACC 199

Qy   191 TGAAGTTGAAAGAAGTTCAAAGATGGGGTTCATATGCCTGTGATATATTTCTGATTCTGATA 250
           |||||| |||||||||||| ||||||| ||||| ||||||||| |||||| ||||||||||
Db   200 CTCGAGCCCGACGAATCCAAAAAGGTCACYCGCCCCATTATGCTGATGTATCTKATCCTGAAT 259

Qy   251 CCGTTCACAAGGTGTTTGCTCAAGTTGCTAAGGATTTTGGTAAGTTGCCATTGCACTTGG 310
           || | || |||||||||||||||||||||| ||||||||| |||||||||||||||||||
Db   260 CTGTCGAAGGCTTGTTTAGCCGAAGTTGTTAAGGAGCACGGGAAAGATCGACA---ACCTGG 316

Qy   311 TTAACACAGCTGGGTTACTGTGGAAAACTTCCCATGTGAAGAGATTACCGACGGTCTCC 370
           |||||||||||||||||||||| ||||| |||||||||||||| ||||| ||||||||
Db   317 TCACCTCAGCTGGCTTCACGGAGAACTTCAGAAGCCGTTAACTACBSCATCGACCGTCC 376

Qy   371 AGAAGGATGGTCGAAGGTTAACTTGTTGGGTTCTTGTAGTGTTTTCAAGCCTTGCTAAGC 430
           |||||||||| ||||||||||||||||||| ||||| |||||||| ||||||| ||
Db   377 GTAAGCTTTGGCGCTGTTAACGTTGACGGTACATATYCTTGCAACATACAGTCGCCAGGY 436

Qy   431 CATTGAT-CAAAGAAGGTATCAAGGTGCTCTTCGTGTTGTTTTGGTTCTATGTGTCTGGT 489
           ||||| |||||||||| |||||||| | | |||| |||||||||||||||||||
Db   437 ACTTTATGCAAAGAAAAGGSTCCGGTGGCATC----GTCATGAITTGTAGCAITGTCCGGA 492

Qy   490 GCCATTGTCAACGAATCCTCAAAACCAAGTTGTCACACGATGTCCAAGGGTGGGTCGTTATC 549
           ||||||||||||||||||||| ||||| |||| ||| || ||||||| |||||||||||||
Db   493 TCCATTGTCAACGTTCCTCAGCCTCAGGCTCCYTATAATCGYGCCAAAGCYGGTGYTGGCGGC 552

Qy   550 CATTTGGGCTAAGACTTTGGGTTGTTGGAANGGGCGTAAGTAGACAACATCAGATTAATTC 609
           ||||| | ||||||| |||||| ||||| ||||| |||| || ||| || || ||||
Db   553 CAACTCGCTCGCTTCYTTGGCCGGTCGAAATGGGCTCAGGCGAAACATCCGAGTTAACTGCATC 612

Qy   610 AACCCAGGTTACATCTCACGGTCCTTTTGACCAAGAANGTTATCAAATGGTAACGAAGA 665
           |||| |||||||||| ||||||||| |||||| ||| |||| || || ||| |||
Db   613 TCTCCSGGTTACATGTTGACTGCACTCACTCAGGAGAAAATCTTGACGACAACACCBGGA 668

Qy   14 TTCCAACTTTTAGATTCGDATGGGCCACTTAACCATTGTCACAGGTGCCCTGTGGGTTTAG 73
           ||||||||||||| |||| ||| |||| | || | ||||||| |||| |||| |||||||
Db   20 TGCCACCTTTACACTCGACGGTCATGTTGGCATCGTTGGCATGGTCCTGTGGTTGG 79

Qy   74 CTGAAGCTTTAATCAAAGGGTTTGTTGGCCTACCGTTCTGACATTGCTTTGCTTGATATCG 133
           | ||| ||| || ||| |||| || | || ||||| ||||| ||| || |||||||
Db   80 GTCTTGTTTATGGGTCAAGGAATGGTATTCTCTGGATCCAACCTTGCTTGGTTCTTGTTGATTATTW 139

RESULT 5

LOCUS   CF701403   836 bp   mRNA   linear   EST 16-AUG-2004
DEFINITION   CCACT16TF C.neoformans strain JEC21 Cryptococcus neoformans var.
               neoformans cDNA clone CCACT16, mRNA sequence.
ACCESSION   CF701403
VERSION   CF701403.1 GI:41555562
KEYWORDS   EST.
SOURCE   Cryptococcus neoformans var. neoformans
  ORGANISM   Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
               var. neoformans)
               Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
               Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
               Filobasidiella.
REFERENCE   1 (bases 1 to 836)
  AUTHORS   Loftus,B.
  TITLE   End sequencing of clones from a full length enriched, normalized
               JEC21 cDNA library
  JOURNAL   Unpublished (2003)
COMMENT   Other_ESTs: CCACT16TR
               Contact: Brendan Loftus
               TIGR
               Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
               Rockville, MD 20850, USA
               Tel: 301-838-3543
               Fax: 301-838-0208

Query Match   13.9%; Score 113.4; DB 7; Length 713;
Best Local Similarity   54.6%; Pred. No. 3.5e-21;
Matches   278; Conservative   0; Mismatches 216; Indels   15; Gaps   2;

Qy   308 TGGTTAACACAGCTGGTTACTGTTGAAAAACTTCCCATGTGAAGATTACCCAGCCAAGAACG 367
              ||| || ||| ||| |||||| | || ||| |||| |||| |||| || ||||| |||||
Db   649 TCGTCACTGCTGCCGGTATTGTCGAAAAACTTTGTCGCTCACGAGTACCCCATCGATAAGA 590

Qy   368 CTGAGAAGAGATGGGTGAAGGTTAACTTGTTGGGGTTTCTTTGTATGTTTCTCAAGGCTTA 427
              ||||| ||| |||| ||| |||||| |||||| |||| |||| |||| ||| || ||| |||
Db   589 TCAAGAAGCTGTTGGACATCAACATTATGGGTACTTGGTTATTCGCACTTGAGGCTGCCA 530

Qy   428 AGCCATTGATCAAAGAAGGTATCAAGGGTGCTTCTTGTTGTTGATTGGTTCTATGTCTG 487
              |||||||||||| |||| ||||||| |||| ||||| ||| |||| |||||||| |||||
Db   529 AGCTTATGCCTGAA--------------GGTGGTTCCATTACCCTCGTCGGCATCTATGAGCG 482

Qy   488 GTGCCATTGTCAACGATCCTCAAAAACCAAGTGTCTACAACAGTCCAAGGGCTGGTGTTA 547
              ||||||||||||||||| |||||| ||||||||| ||||||| ||||||||||| ||||
Db   481 GTAGCATTGTCAACGGTTCCTCAAACCTACAACTTTTTCCAAGGCTGCTGCTGTGC 422

Qy   548 TCCATTTGGCTAAGACTTTGGCTTGTGAATGGGCTAGTACAACATCAGAGTTAATTCTT 607
              ||||||||||| ||||||||| ||||| |||||| ||| ||||||||||||| ||||||
Db   548 TCAAGAAGCTGTTGGACATCAACATTATGGGTACTTGGTTATTCGCACTTGAGGCTGCCA 546

Qy   608 TAAACCCAGGTTACATCTACGGGTCCTTTGACCAAAGAATGTTATCAATGGTAACGAAGAAT 667
              ||||||||||| ||| |||||| ||||||||||||||| ||||||||||| || |||||||
Db   421 GACACATGGCTCAGATCCCTCGCCGTCGAATGGGCTCTCAAAGGGTATCCGTGTCAAGCGTC 362

Qy   728 TTGGTGCTGCTGTTTTGTACTTGCTTTCTGAATCTGCTTCATCACTACTGGTGCCAGCT 787
              ||||| | ||||||| ||||||||||| |||| |||| || |||| || |||||||||||
Db   241 AGGGTGCCGGTCATTACCTTGCTTCTGA---CAGCTCCAAGTACACCACTGGTGCTGAGA 185

Qy   788 TACTGGTTGATGGTGGTTTGGGTTTCACTTCTTCG 816
              ||| ||||||||| ||||| ||||||||||||| |
Db   184 TCATGATTGACGCGGCGGTTACACTTGCTTG 156

RESULT 6

LOCUS   CF715948   836 bp   mRNA   linear   EST 16-AUG-2004
DEFINITION   CCAB241TF C.neoformans strain JEC21 Cryptococcus neoformans var.
               neoformans cDNA clone CCAB241, mRNA sequence.
ACCESSION   CF715948
VERSION   CF715948.1 GI:41570107
KEYWORDS   EST.
SOURCE   Cryptococcus neoformans var. neoformans
  ORGANISM   Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
               var. neoformans)
               Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
               Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
               Filobasidiella.
REFERENCE   1 (bases 1 to 836)
  AUTHORS   Loftus,B.

ORIGIN
FEATURES   Location/Qualifiers
     source   1..836
                     /organism="Cryptococcus neoformans var. neoformans"
                     /mol_type="mRNA"
                     /strain="JEC21"
                     /db_xref="taxon:40410"
                     /clone="CCAB241"
                     /clone_lib="C.neoformans strain JEC21"
                     /note="Vector: pCMVSport6; Site 1: NotI EcoRV; The full
                     length, normalized library was prepared from a variety of
                     conditions using RNA provided by Joseph Heitman and
                     Jennifer Lodge"
Seq primer: TF.
Email: crypt@tigr.org

Query Match   13.9%; Score 113.4; DB 7; Length 836;
Best Local Similarity   54.6%; Pred. No. 3.7e-21;
Matches   278; Conservative   0; Mismatches 216; Indels   15; Gaps   2;

Qy   308 TGGTTAACACAGCTGGTTACTGTTGAAAAACTTCCCATGTGAAGATTACCCAGCCAAGAACG 367
              ||| || ||| ||| |||||| | || ||| |||| |||| |||| || ||||| |||||
Db   665 TCGTCACTGCTGCCGGTATTGTCGAAAAACTTTGTCGCTCACGAGTACCCCATCGATAAGA 606

Qy   368 CTGAGAAGAGATGGGTGAAGGTTAACTTGTTGGGGTTTCTTTGTATGTTTCTCAAGGCTTA 427
              ||||| ||| |||| ||| |||||| |||||| |||| |||| |||| ||| || ||| |||
Db   605 TCAAGAAGCTGTTGGACATCAACATTATGGGTACTTGGTTATTCGCACTTGAGGCTGCCA 546

Qy   428 AGCCATTGATCAAAGAAGGTATCAAGGGTGCTTCTTGTTGTTGATTGGTTCTATGTCTG 487
              |||||||||||| |||| ||||||| |||| ||||| ||| |||| |||||||| |||||
Db   545 AGCTTATGCCTGAA--------------GGTGGTTCCATTACCCTCGTCGGCATCTATGAGCG 498

Qy   488 GTGCCATTGTCAACGATCCTCAAAAACCAAGTGTCTACAACAGTCCAAGGGCTGGTGTTA 547
              ||||||||||||||||| |||||| ||||||||| ||||||| ||||||||||| ||||
Db   497 GTAGCATTGTCAACGGTTCCTCAAACCTACAACTTTTTCCAAGGCTGCTGCTGTGC 438

Qy   548 TCCATTTGGCTAAGACTTTGGCTTGTGAATGGGCTAGTACAACATCAGAGTTAATTCTT 607
              ||||||||||| ||||||||| ||||| |||||| ||| ||||||||||||| ||||||
Db   437 GACACATGGCTCAGATCCCTCGCCGTCGAATGGGCTCTCAAAGGGTATCCGTGTCAAGCGTC 378

Qy   608 TAAACCCAGGTTACATCTACGGGTCCTTTGACCAAAGAATGTTATCAATGGTAACGAAGAAT 667
              ||||||||||| ||| |||||| ||||||||||||||| ||||||||||| || |||||||
Db   377 TTAGTCCGGGTTACGTCCTCACCAAACTTGACTAAGGTCATTCTCGACGCGCCAACCCCGTTC 318

Qy   668 TGTACAACAGATGGATCTCTGGTATCCCACACAAACAAGAAGAAATGTCCGGACCAAAAGGAGTTA 727
              ||| ||||| ||||| |||||| ||||| || ||||||| ||||||| ||| |||||| |||||
Db   317 TCCGTGACGAGTGGCTCAACCGTATCCCCATGGGTCGAATGGGCCATCTATGAGCG 258

Qy   728 TTGGTGCTGCTGTTTTGTACTTGCTTTCTGAATCTGCTTCATCACTACTGGTGCCAGCT 787
              ||||| | ||||||| ||||||||||| |||| |||| || |||| || |||||||||||
Db   257 AGGGTGCCGGTCATTACCTTGCTTCTGA---CAGCTCCAAGTACACCACTGGTGCTGAGA 201

Qy   788 TACTGGTTGATGGTGGTTTGGGTTTCACTTCTTCG 816
              ||| ||||||||| ||||| ||||||||||||| |
Db   200 TCATGATTGACGCGGCGGTTACACTTGCTTG 172

REFERENCE   1 (bases 1 to 836)
  AUTHORS   Loftus,B.

RESULT 7

LOCUS   CF817896   777 bp   mRNA   linear   EST 01-APR-2004
DEFINITION   EST695278 Coccidioides posadasii saprobic phase cDNA library, size
               greater than 4kb Coccidioides posadasii cDNA clone CIBB147 5' end,
               mRNA sequence.
ACCESSION   CF817896
VERSION   CF817896.1 GI:45923774
KEYWORDS   EST.
SOURCE   Coccidioides posadasii
  ORGANISM   Coccidioides posadasii
               Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
               Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1 (bases 1 to 777)
  AUTHORS   Gardner,M.J. and Cole,G.T.
  TITLE   Analysis of gene expression in Coccidioides posadasii expressed sequence tags
  JOURNAL   Unpublished (2003)
  Other ESTs: EST695277
COMMENT   Contact: Gardner MJ
               The Institute for Genomic Research
               9712 Medical Center Drive, Rockville, MD 20850, USA
               Tel: 301 838 3519
               Fax: 301 838 0208
               Email: gardner@tigr.org
               Seq primer: M13 Reverse.
FEATURES   Location/Qualifiers
     source   1..777
                     /organism="Coccidioides posadasii"
                     /clone="CIBB147"
                     /mol_type="mRNA"
                     /strain="C735"
                     /db_xref="taxon:199306"
                     /dev_stage="saprobic phase (mycelia)"
                     /lab_host="E. coli DH10B, T1 phage resistant"
                     /clone_lib="Coccidioides posadasii saprobic phase cDNA
                     library, greater than 4kb"
                     /note="Vector: pExpress 1; Site 1: Not I; Site 2: Eco RV;
                     Coccidioides posadasii saprobic phase cDNA library, size
                     fractionated cDNA > 4 kb"

ORIGIN

Query Match   13.8%; Score 112.8; DB 7; Length 777;
Best Local Similarity   54.1%; Conservative 0; Mismatches 227; Indels 6; Gaps 2;

Matches 275;

Qy   304 CACTTGGTTAACACAGCCTGGTTACTGTGAAAAACTTCCCATGTGAAGAAGAACG 363
          ||||| ||| ||| || ||||||||| ||| | ||||||| ||||| |||
Db   708 CATCTGGTTCACATCTGCGGGGGTTCACTGAAAAACTTTGACGCGCCTACCAGCCAAG 649

Qy   364 AACGCTGAGAAGATGGTGAAGGTTAACTTGTTGGTGTTCTTTGTATGTTTCTCAAGCCTTT 423
          ||||||||||||||||| ||||| |||||||||||| | | ||||| |||||| |
Db   708 CGCATGCAGAAAACTCTGGGTCCGTCAACGTTGATGGGTCAATATCTTTTGTGCAGTTGCTGTT 649

Qy   424 GCTAAGCCATTGATCAAGAGAAGGTATCAAGGGGTGCTTCTCGTGTTGTTTGATTGGTTCTATG 483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   648 CGCATGCAGAAAACTCTGGGTCCGTCAACGTTGATGGGTCAATATCTTTTGTGCAGTTGCTGTT 589

Qy   484 TCTGGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTCTACAACAGTCCAAGGCTGCG 543
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   588 GCCAAACATCTGAT---GGCTCGCAAGTCGCCTGGCAGCATTGTGGTTCATTGGAAGCATG 532

Qy   544 GTTATCCATTGGCTAAGACTTTGGCTTGTGAAATGGGCTAAGTACAACATCAGAGTTAAT 603
          |||||||||||||||||||| || |||||||||||||||||||||||||||||||||
Db   531 TCTGGTGCCATTGTTAACGTGCCACAACATGTCAAGGCGGGAAATGCGGAGAGTGAAC 472

Qy   604 TCTTTAAACCCAGGTTACATCTACGGGTCCTTTGACCAAGAATGTTATCAATGGTTACGAA 663
          | ||| || ||| |||||||||||||||||||||||||||||||||||||||||||
Db   471 ATAAGGCACTTGGCTGCTCGTTGCCTCTCCTTGGCAGTTGAAATGGGACGTGGGAAATCAGAAG 412

Qy   664 GAATTGTACAACAGATGGAGATCTCGGTATCCCACAACAACAAAGAGAATGTCCGAACAAAGGAA 723
          ||||||||||||||||||||||||||||||||||||||| |||| |||| |
Db   411 TGTATCAGTCAGGATTACATGTTGACTGCCCTGACCCGCAAGATTCTTGACGACAACCT 352

Qy   724 TACATTGGGTCGTGTTTTGTTGTTGTACTTGCTTCTTCATCACACTACTGGTGCC 783
          |||| |||| || |||| || ||| |||||| |||||||||||||||||||
Db   351 GATCTCAAAGAGAGAAGTGGACCTCCCTCATTCCACACGACAAAATGGGTCACACTGAAGAC 292

Qy   784 AGCTTACTGGTTGATGGGGGTTTCACTT 811
          |||||||||||||||||||||||||||
Db   291 CTCATGGGGCCCAGTGACGTTCTTATTGAGTGA---TGCGAGCAAAATATGTTACTGGGGCA 235

RESULT 7

LOCUS   CF817896/c   777 bp   mRNA   linear   EST 01-APR-2004
DEFINITION   EST695278 Coccidioides posadasii saprobic phase cDNA library,
               greater than 4kb Coccidioides posadasii cDNA clone CIBB147 5' end,
               mRNA sequence.

TITLE   End sequencing of clones from a full length enriched, normalized
               JEC21 cDNA library
JOURNAL   Unpublished (2003)
               Other ESTs: CCAB241TR
COMMENT   Contact: Brendan Loftus
               TIGR
               Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
               Rockville, MD 20850, USA
               Tel: 301-838-3543
               Fax: 301-838-0208
               Email: crypt@tigr.org
               Seq primer: TF.
FEATURES   Location/Qualifiers
     source   1..836
                     /organism="Cryptococcus neoformans var. neoformans"
                     /mol_type="mRNA"
                     /strain="JEC21"
                     /db_xref="taxon:40410".
                     /clone="CCAB241"
                     /clone_lib="C.neoformans strain JEC21"
                     /note="Vector: pCMVSport6; Site 1: NotI EcoRV; The full
                     length, normalized library was prepared from a variety of
                     conditions using RNA provided by Joseph Heitman and
                     Jennifer Lodge"

ORIGIN

Query Match   13.9%; Score 113.4; DB 7; Length 836;
Best Local Similarity   54.6%; Conservative 0; Mismatches 216; Indels 15; Gaps 2;

Matches 278;

Qy   308 TGGTTAACACAGCTGGTTACTGTGAAAAACTTCCACATGTGAAGAAGAACG 367
          |||||||||| || ||||||||||||||||||| | |||||||||||||
Db   670 TCGTCACTCGCTGCCGGTATTGTCGCACTTGTGCCCTCACGAGTACCCCATCGNTAAGA 611

Qy   368 CTGAGAAGAATGGTGAAGGTTAACTTGTTGGTGTTTGTTGTATGTTTCTCAAGCCTTTGCTA 427
          | |||||| || |||||||||||||||||||||| || || ||| ||||| |
Db   610 TCAAGAAGCTGTTCGGACATCAACATTATAGGGTACTTGGGTTATTGCGGCACTTGAGCGCTGCCA 551

Qy   428 AGCCATTGATCAAAGAAGGTATCAAGGGGTGCCTTCTGTTGTTTTGATTGGTTCTATGTCTG 487
          |||||||||||||| --------GGTGGGTTCATTACCCTGTTCGCATCTATGAGCG 503
Db   550 AGCTTATGCCTGAA----------------GGTGGGTTCATTACCCTGTTCGCATCTATGAGCG 503

Qy   488 GTGCCATTGTCAACGATCCTCAAAAACCAAGTGTTCTACAACAGTCCAAGGCTGGGTGTTA 547
          |||||||||||||||||||||||||||||| || ||||||| || || |||||||||||
Db   502 GTAGCCATTGTCAACGTTCCTCAACCTACAACATTTTCCAAGGCTGCTGCTGC 443

Qy   548 TCCATTTGGCTAAGACTTTGGCTTGTGTGATGGGCTAAGTACAACATCAGAGTTAATTCTT 607
          |||||||||||||||||||||| || ||||||||||||||||||||||||||||||||
Db   442 GACACATGGCTCGATCCCTCGCCGTGAATGGGGTCTCAAGGTATCCGTGTCAAACGCTC 383

Qy   608 TAAACCCAGGTTACATCTACGGGTCCTTTGACCAAGAATGTTATTCAAATGGTAACGAAGAT 667
          || ||| ||||||||||||||||||||||||||||||||| |||||||||||||
Db   382 TTAGTCCGGGTTACGTCTCACCAACTTGACTAGGGTCAATTGCAACCCGGTC 323

Qy   668 TGTACAACAGGATGGATCTCTGGGTATCCCACAACAAAGAGAATGAGAATACA 727
          |||||||||| ||| ||||| |||||||||||||||||||||||||||||
Db   322 TCCCTGACCGAGTGGCGTACCACCGTATCCCCATGGGTCCGAATGGCCACCTCTGATCTCA 263

Qy   728 TTGGTGCTGCTGTTTTGTACTTGCTTTCTTCATACACTACTGGTGCCCAGCT 787
          |||||||||||| | ||||||||||||||| ||||| ||||||| ||
Db   262 AGGGTGCCGTCATTACCGTTGGTGTTTCACTTGTTCTGACAGCGTCCTGAGA 206

Qy   788 TACTGGTTGATGGTGGTTTCACTTCTTGG 816
          ||||| |||||| || ||||||||||||
Db   205 TCATGATTGACGGGGGTTACACTTGCTTG 177

```
Db    234 GATTTGAGGGTTGATGGCGGGGTACACTT 207
          |||||||||||||||||||||||||||||
```

RESULT 8

CF824542/c

```
LOCUS       CF824542    787 bp    mRNA    linear    EST 01-APR-2004
DEFINITION  EST701924 Coccidioides posadasii saprobic phase cDNA library, 2 to
            4 kb Coccidioides posadasii cDNA clone CIDAR90 3' end, mRNA
            sequence.
ACCESSION   CF824542
VERSION     CF824542.1  GI:45930599
KEYWORDS    EST.
SOURCE      Coccidioides posadasii
  ORGANISM  Coccidioides posadasii
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1  (bases 1 to 787)
  AUTHORS   Gardner,M.J. and Cole,G.T.
  TITLE     Analysis of gene expression in Coccidioides posadasii mycelia and
            spherules via expressed sequence tags
  JOURNAL   Unpublished (2003)
COMMENT     Other ESTs: EST701925
            Contact: Gardner MJ
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: gardner@tigr.org.
FEATURES             Location/Qualifiers
     source          1..787
                     /organism="Coccidioides posadasii"
                     /mol_type="mRNA"
                     /strain="C735"
                     /db_xref="taxon:199306"
                     /clone="CIDAR90"
                     /dev_stage="saprobic phase (mycelia)"
                     /lab_host="E. coli DH10B, T1 phage resistant"
                     /clone_lib="Coccidioides posadasii saprobic phase cDNA
                     library, 2 to 4 kb"
                     /note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
                     Coccidioides posadasii saprobic phase cDNA library, size
                     fractionated cDNA 2 to 4 kb"
ORIGIN
```

```
Query Match       13.8%;  Score 112.8;  DB 7;  Length 787;
Best Local Similarity   54.1%;
Matches  275;  Conservative   0;  Mismatches 227;  Indels   6;  Gaps   2;
```

```
Qy    304 CACTTGGTTAACACAGCTGGTTACTGTGTGAAAGATTACCCAGCCAAG 363
          |||| |||||||||||||| ||||||||||||||||||||||||||||
Db    706 CACTCGGTCACATCTGAT---GGCTCGGCAAGTCGCCTGGGGGTTCACTGAAAAACTTTGACGCCGTGTCCTACCCGCATGAT 647

Qy    364 AACGCTGAGAAGATGGTGAAGGTTAACTTGTTGGGTTGTTCAAGCCTTT 423
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    646 CCGATGCAGAAAACTCTGGTCCGTCACGTTGATGGGTCAATATCTTTT 587

Qy    424 GCTAAGCCATTGATCAAAGAAGGTATCAAGGGTGCTTCTCTGTTTCTATG 483
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 GCCAAACATCTGAT---GGCTCGGCAAGTCGCCTGGCCAGCATTGTTGTTCATTGGAAGCATG 530

Qy    484 TCTGGGTGCCATTGTTCAACGATCCTCAACAAGTTGTCTACAACACATC 543
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 TCTGGGTGCCATTGTTAACGATGGTCTCAGCCACAAGTCTTCAACGTTGCG 470

Qy    544 GTTATCCATTGGCTAAGACTTGGCCTTGTTCGGAATGGGGTTAAGTACAACATCAGAGTTAAT 603
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 ATAAGGGACACTTGGCTGCTTCCTTGGCACGTTGAAATGGGAATCAGAGTGAAC 410

Qy    604 TCTTTAAACCCAGGTTACATTACGGTTCTTGGCCAGGTTGAATGGGA 663
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    409 TGTATCAGTCAACCAGGATACAATGTTGACTGCCCCTGACCCGGAAGAATTCTTGACGACAACCCT 350
```

RESULT 9

CF812382/c

```
LOCUS       CF812382    738 bp    mRNA    linear    EST 01-APR-2004
DEFINITION  EST689764 Coccidioides posadasii saprobic phase cDNA library,
            greater than 4kb Coccidioides posadasii cDNA clone CIBA394 3' end,
            mRNA sequence.
ACCESSION   CF812382
VERSION     CF812382.1  GI:45918260
KEYWORDS    EST.
SOURCE      Coccidioides posadasii
  ORGANISM  Coccidioides posadasii
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1  (bases 1 to 738)
  AUTHORS   Gardner,M.J. and Cole,G.T.
  TITLE     Analysis of gene expression in Coccidioides posadasii mycelia and
            spherules via expressed sequence tags
  JOURNAL   Unpublished (2003)
COMMENT     Other ESTs: EST689765
            Contact: Gardner MJ
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: gardner@tigr.org.
FEATURES             Location/Qualifiers
     source          1..738
                     /organism="Coccidioides posadasii"
                     /mol_type="mRNA"
                     /strain="C735"
                     /db_xref="taxon:199306"
                     /clone="CIBA394"
                     /dev_stage="saprobic phase (mycelia)"
                     /lab_host="E. coli DH10B, T1 phage resistant"
                     /clone_lib="Coccidioides posadasii saprobic phase cDNA
                     library, greater than 4kb"
                     /note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
                     Coccidioides posadasii saprobic phase cDNA library, size
                     fractionated cDNA > 4 kb"
ORIGIN
```

```
Query Match       13.6%;  Score 111.2;  DB 7;  Length 738;
Best Local Similarity   53.9%;
Matches  274;  Conservative   0;  Mismatches 228;  Indels   6;  Gaps   2;
```

```
Db    232 GATTTGAGGGTTGATGGCGGGGTACACTT 205

Qy    664 GAATTGTACAACAGAGATGGATCTCTGGGTATCCCACTAACAACAAAGAATGTCGGACAAAAAGGAA 723
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 GATCTCAAAGAGAGAAGAGTGGGACCTCCCTCATTCCACAAGGCAAAAATGGGTACAACTACTTGGTGCC 783

Qy    724 TACATTGGTTGCTCGTTTTTGTACTTGCTTTCTGAAATCTGCTTCATACACTACTTGGGGCA 783
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 CTCATGGGCGCCCAGTGACGTTCTTTATTGAGTGA---TGCCAGCAAAATATGTTACTGGGGCA 233

Qy    784 AGCTTACTGCGTTGATGGTGGGTTCACTT 811
          |||||||||||||||||||||||||||||
Db    232 GATTTGAGGGTTGATGGCGGGGTACACTT 205

Qy    304 CACTTGGTTAACACAGCTGGTTACTGTGTGAAAGATTACCCAGCCAAG 363
          |||| |||||||||||||| ||||||||||||||||||||||||||||
Db    612 CACTCGGTCACATCTGAT---GGCTCGGCAAGTCGCCTGGGGGTTCACTGAAAAACTTTGACGCCGTGTCCTACCCGCATGAT 553

Qy    364 AACGCTGAGAAGATGGTGAAGGTTAACTTGTTGGGTTGTTCAAGCCTTT 423
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 CGGATGCAGAAAACTCTGGTCCGTCACGTTGATGGGTCAATATCTTTTGTTTTTGATTGGTTGCTGTT 493

Qy    424 GCTAAGCCATTGATCAAAGAAGGTATCAAGGGTGCTTCTCTGTTTCTATG 483
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 GCCAAACATCTGAT---GGCTCGGCAAGTCGCCTGGCCAGCATTGTTGTTCATTGGAAGCATG 436

Qy    484 TCTGGGTGCCATTGTTCAACGATCCTCAAAAACCAAGTTGTCTACAACACATCGCAAGGCTGGT 543
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 TCTGGGTGCCATTGTTAACGATGGTCTCAGCCACAAGTCTTCAACGTTGCG 376
```

RESULT 10

```
LOCUS       CF824543                 676 bp    mRNA    linear   EST 01-APR-2004
DEFINITION  EST701925 Coccidioides posadasii saprobic phase cDNA library, 2 to
            4 kb Coccidioides posadasii cDNA clone CIDAR90 5' end, mRNA
            sequence.
ACCESSION   CF824543
VERSION     CF824543.1  GI:45930600
KEYWORDS    EST.
SOURCE      Coccidioides posadasii
  ORGANISM  Coccidioides posadasii
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1  (bases 1 to 676)
  AUTHORS   Gardner,M.J. and Cole,G.T.
  TITLE     Analysis of gene expression in Coccidioides posadasii mycelia and
            spherules via expressed sequence tags
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: EST701924
            Contact: Gardner MJ
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: gardner@tigr.org
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1..676
                     /organism="Coccidioides posadasii"
                     /mol_type="mRNA"
                     /strain="C735"
                     /clone="CIDAR90"
                     /db_xref="taxon:199306"
                     /clone_lib="Coccidioides posadasii saprobic phase cDNA
                     library, 2 to 4 kb"
                     /dev_stage="saprobic phase (mycelia)"
                     /lab_host="E. coli DH10B, T1 phage resistant"
                     /note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
                     Coccidioides posadasii saprobic_phase cDNA library, size
                     fractionated cDNA 2 to 4 kb"
ORIGIN
```

Query Match 13.0%; Score 106.4; DB 7; Length 676;
Best Local Similarity 58.6%;
Matches 204; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

```
Qy    544 GTTATCCATTTGGCTAAGACTTTGGCTTGTGAATGGGTAAGTACAACATCAGAGGTTAAT 603
          |||| ||||| ||| ||  |  |||||| |||||||||| | | ||| ||| |||||||||
Db    375 ATAAGGCACTTGCGTCGTTCCTTGGCGTTGAATGGGGCAAGTGTGGGAATCAGAGTGAAC 316

Qy    604 TCTTTAAACCAGGTTACATCTACAGGGTCGTTTGACCAACAGAGTATGGTTATCAAATGGT 663
          ||| | |  |||||| || | || |||| || |||| || |||| ||| || |||  | ||
Db    315 TGTATCAGTTCAGGGATACATGTTGACTGCCTGAAGATTCTTGACGACAACCCT 256

Qy    664 GAATTGTACAACAGATGAGATCTCTGGTATCCCACAACAAAGAAATGTCCGAACAAAGGAA 723
          || |||| ||  || |||| ||||| ||| | |||| | || | || || ||| ||||| |
Db    255 GACTCAAAGAAGATGGGACTCCCTCATTCCACAAGGCCAAAAATGGGTTACACCTGAAGAC 196

Qy    724 TACATTGGTACGTGGTGCGTCTTTGTTTTGTTTACTCGTATCAAATCTGATGGTTCTGCTG 783
          || | |||| || | |  | ||| | | ||  |||| |||||| | |||| |||| | || |
Db    195 CTCAGGGACCCCAGTGACGTTCTTATTGAGTGA---TGCGAGCAAAATATGTTACTACTGG 139

Qy    784 AGCTTACTGGTTGATGGTGCGTTTCACTT 811
          | | |||||||||| |||||||| |||||
Db    138 GATTTGAGGGTTGAATGGCGGGTACACTT 111
```

RESULT 11

```
LOCUS       AJ638396                 718 bp    mRNA    linear   EST 05-MAY-2004
DEFINITION  AJ638396 MgC Mycosphaerella graminicola cDNA clone mgc06e02f, mRNA
            sequence.
ACCESSION   AJ638396
VERSION     AJ638396.1  GI:47031453
KEYWORDS    EST.
SOURCE      Mycosphaerella graminicola
  ORGANISM  Mycosphaerella graminicola
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
            Chaetothyriomycetes incertae sedis; Mycosphaerellaceae;
            Mycosphaerella.
REFERENCE   1  (bases 1 to 718)
  AUTHORS   Keon,J.P.R., Hargreaves,J.A., Antoniw,J.F. and Hammond-Kosack,K.
  TITLE     Analysis of expressed sequence tags from the wheat fungal leaf
            blotch pathogen, Mycosphaerella graminicola (anamorph Septoria
            tritici)
  JOURNAL   Fungal Genet. Biol. (2004) In press
COMMENT     Contact: Keon J
            Plant Pathogen Interactions Division,
            Rothamsted Research,
            Harpenden, Herts, UNITED KINGDOM
            Tel: +44 (0)1582 763133
            Fax: +44 (0)1582 760981
            Email: john.keon@bbsrc.ac.uk
            Insert Length: 800  Std Error: 100.00
            Seq primer: M13 reverse.
FEATURES             Location/Qualifiers
     source          1..718
                     /organism="Mycosphaerella graminicola"
                     /mol_type="mRNA"
                     /strain="Strit"
                     /db_xref="taxon:54734"
                     /clone="mgc06e02f"
                     /clone_lib="MgC"
                     /note="Vector: pSPORT1; Library constructed from senescent
                     wheat leaves 21-25 days after infection with
                     Mycosphaerella graminicola exhibiting abundant hyphal
                     growth and asexual sporulation"
ORIGIN
```

Query Match 12.6%; Score 102.8; DB 1; Length 718;
Best Local Similarity 53.0%;
Matches 268; Conservative 0; Mismatches 232; Indels 6; Gaps 2;

```
Db    181 ACAATGCTGCAAAGGCTGCGGATAAGGCACTTGGCTTCCTTGGCGTTGACGTTGAAATGGGGCAA 240
          |||||||                | || ||  ||||||||| || |||||| ||||||  || |
Qy    524 ACAACATGTCCAAGGCTCGGTGTGTTATCCATTTGGCTAAGACTTTGGCTTGTGAATGGGTA 583

Db    584 AGTACAACATCAGAGGTTAATTCTTTAAACCAGGTTACATCTACAGGGTCGTTTGACCAAG 643
          ||||| ||| |  |||| || ||| | |  |||| ||  |  ||| |||| | |||| ||
Qy    464 TTGTTTTTGATTGGTTCTAATGTCTGGTGCATCCTCAAACATCCAAAACCAGTTGTCT 523

Db    241 GTGTGGGAATCAGAGTGAACTGTATCAGTTCAGGGATACATGTTGACTGCCTGAAGATTCTTGACGACAACCCT 300
          |||| |  || |||| || ||||| || ||| ||   || ||||| | |||| || | ||  ||| ||   |
Qy    121 TTGTGGTCATTGGAAGCATGTCTCGGTGCCATTGTTAACGTGCCTCAGCCACCAACAAAGCGCCGT 180

Db    644 ATGTTATCAAATGGTGAAGGTAAGCAAACCCTGATCTCAAGAGATGGAATTGTACAACAGATGTCCAACAAA 703
          | | |||||  || |||||    || ||  |||  ||||   ||| |  ||| |||||||  || || ||  |
Qy    364 AACGCTGAGAGAGATGGTGAAGGTTAACGTCGTTGGGGTTTCTTGTATGTTTCTTGTCAAGCCTTT 423

Db    301 AGATTCTTGACGACAACCCTGATCTCAAAGAAGATGGGACTCCCTCATTCCACAAGGCA 360
          ||||| ||||||| ||||  ||| | ||| | ||| | || || |||||||||||
Qy    304 CACTTGGTTAACACAGCTGGTTACTGTGAAAAACTTCCCATGTGAAGATTACCACGCCAG 363

Db    704 GAATGTCCGAACCAAAGGAATACATTGGTACGTGGTGCGTCTTTGTTTTGTTTACTCGTATCAAATCTGCTG 763
          || ||| ||| || ||| |||| | |||| ||| | | || || | | ||| |||| |||| ||  | || |
Qy    49  CACCTCGTCACTTCCGCTGCTTCACCGAGAACTTCGACCGCTACGAGCCACACAGAC 108

Db    361 AAAATGGGTACACCTGAAGACCTCATGGGACCCCAGTGACGTTCTTATTGAGTGA---TGCGA 417

Db    764 CTTCATACACTACTGGGTGTGCCAGCTTACTGGTTGATGGTGCGTTTCTTATTGAGTGA---TGCGA 811
          ||||| ||  ||||||||| || ||| ||||||||| |||||||| || |||| |||| ||  |  |
Qy    109 CGCATGCAGAAGACTTTGGGAGAGTCAACGTCGACGGCACATACACTGGTCTGCACCGTG 168

Db    418 GCAAATATGTTACTGGGGCAGATTTGAGGGTTGAATGGCGGGTACACTT 465
```

RESULT 12
CO136295
LOCUS   CO136295   615 bp   mRNA   linear   EST 17-JUN-2004
DEFINITION   EST830966 Aspergillus flavus Normalized cDNA Expression Library
               Aspergillus flavus cDNA clone NAFBQ16 5' end similar to
               D-arabinitol 2-dehydrogenase [ribulose forming] (EC 1.1.1.250)
               (ARDH). [Yeast] (Candida tropicalis], mRNA sequence.
ACCESSION   CO136295
VERSION   CO136295.1   GI:48885273
KEYWORDS   EST.
SOURCE   Aspergillus flavus
ORGANISM   Aspergillus flavus
               Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
               Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE   1 (bases 1 to 615)
AUTHORS   Yu,J., Whitelaw,C.A., Nierman,W.C., Bhatnagar,D. and Cleveland,T.E.
TITLE   Aspergillus flavus expressed sequence tags for identification of
               genes with putative roles in aflatoxin contamination of crops
JOURNAL   FEMS Microbiol. Lett. (2004) In press
COMMENT   Contact: Yu J
               Food and Feed Safety Research Unit
               USDA/ARS, Southern Regional Research Center
               1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA
               Tel: 504 286 4405
               Fax: 504 286 4419
               Email: jiuyu@srrc.ars.usda.gov
               Contact Dr. Yu at USDA/ARS SRRC (jiuyu@srrc.ars.usda.gov) for clone
               information
               PCR Primers
               FORWARD: M13F
               BACKWARD: M13R
               Seq primer: M13 Forward
               POLYA=No

FEATURES   Location/Qualifiers
     source   1..615
                     /organism="Aspergillus flavus"
                     /mol_type="mRNA"
                     /strain="NRRL 3357"
                     /db_xref="taxon:5059"
                     /clone="NAFBQ16"
                     /sex="asexual mycelia"
                     /cell_type="mycelia"
                     /dev_stage="developmental stages from 18 to 96 hours"

Qy   424 GCTAAGCCATTGATCATCAAAGAAGAGTATCAAGGGTGCTTCTGTTGTTGTTTTGATTGGTTCTATG 483
          ||||| | || |||| |||||||||||||||||||||||| ||||| |||| ||
Db   169 GCCAAGC---ATCTCATGGAGCCAAGGCTTCTGGCACGTGGTGATGATTGGTTAGCATG 225

Qy   484 TCTGGGTGCCATTGTCAAGGATCCTCAAAACCAAGTTGTCTACACACATGTCCAAGGCTGGT 543
          |||||| |||||||||| || ||||| | || |||| | | || ||||| |||||||| |
Db   226 TCTGGGCGCCATTGTCAATGTGCCTCAGCCACCAGGCTCCCTCACAACGCGGCCAAGGCTGCT 285

Qy   544 GTTATCCATTTGGCTAAGACTTTGGCTTGTTGTGAATGGGCTAAGTACTACAACATCAGAGTTAAT 603
          || | | || ||| | ||| | |||| | || |||||| | |||| | ||| ||||| ||
Db   286 GTCCGGCCCATCCGCCAGCACTCGCGCTCGCCAGTCGAGTGGGGCTCACGCTGGCCATTCCGTGTCAAC 345

Qy   604 TCTTTAAACCCAGGTTACATCTCTACACGGTCCCTTGACCAAGAATGTTATCAATGTTAACGAA 663
          | | | ||||| | |||| | | | | || || |||||| | || | | ||| ||
Db   346 TGCATCTCTCCTGGCACATGTCACTGCCACTACCAAGAAGAATCCTTGACGACACCACCCC 405

Qy   664 GAATTGTACACACAGATAGGATCTCTGGTATCCCACACAACAAAGAATGTCCGGACCAAAGGAA 723
          || | | |||| | | | | || | ||||| |||| | | ||| | ||| || |||| |
Db   406 GAGCTCCGCCAACAAGTGACCAACCTCATCCACACAGGGCAGTGGGGTCGCCCAGAAGAC 465

Qy   724 TACATTGGTGCTGCTGTTTTCTGCTTTCGGAATCTGCTGCTTCATACTACTACTGGTGCC 783
          || | |||||| | ||||| | || ||| | |||| | | || |||||||||||||
Db   466 CTGATGGGCGGCCCGTGACCTTCCTCCTTCCAGGGA---CGCCGCCGGATAGGTGACCGGGCGCC 522

Qy   784 AGCTTACTTGGTTCATTGGTTGGTTTTCAC 809
          |||||| ||| || ||| |||| || |
Db   523 GATCTCCGTGTTGACGGTGGATACAC 548

ORIGIN
                     /lab host="E. coli DH10B T1 resistant cells"
                     /clone_lib="Aspergillus flavus Normalized cDNA Expression
                     Library"
                     /note="Vector: pBluescript (SK+) (Stratagene), antibiotic
                     selection marker: Carbenicillin; Site 1: NotI, at the 5
                     prime end; Site_2: EcoRI, at the 3 prime end; This
                     normalized cDNA expression library was constructed using a
                     mixture of mycelial cells grown under eight different
                     medium conditions and harvested at 5 time points (18, 24,
                     48, 72, 96 hours). The poly-A sequence was trimmed off
                     before ligating to vector."

Query Match   12.4%;   Score 100.8;   DB 7;   Length 615;
Best Local Similarity   51.2%;   Pred. No.1.4e-17;
Matches   287;   Conservative   0;   Mismatches 267;   Indels   6;   Gaps   2;

Qy   196 TTGAAAGAAGTTCCAAAAGATGGGTTCATTAGCCTGTGTGATTATTTCGTCGTGACATGG 255
          ||| | | | || | | ||| |||||| ||||| | | || ||||| |||
Db   38 TTGGAACAAAATGCCCAAGGTCACCGCCGCCCACTACGGCTAAGTTTTGATTTCTGATTACCGTT 97

Qy   256 CACAAGGTGTTTGCTCAAGTTGTCTACACACATGTCCAAGGCTGGTGTTACCCATTG 315
          |||| || | | | | || | | | | ||||| |||| || | | || || |
Db   98 AACGATGCCCTCTCCGGATTATTATCTCCAAGCACACGGCAAGAT---CGACAAACCTGGTCACC 154

Qy   316 ACAGCTGGTTACTGTGAAAAACTCCCAATGTGAAGATTACCAAGCGATGAGAAG 375
          ||| || | || |||||||||| | | ||| |||| |||| || ||| |||
Db   155 TCCGCCGGATTCACGGAAAACTTCGATGGTCCCATCCTACCGACGCTGCAAAAG 214

Qy   376 ATGGTGAAGGTTAAACTTGTTGGGTTCTTTGTATGGTTGCTAAGGCCATTG 435
          || |||||| |||| | ||| || || || | | ||||||||| |
Db   215 CTTTGGGGCGGTTAACGTGAATGGAACATACCTTTCGCCACCGTGTCGCCAAGC---AC 271

Qy   436 ATCAAAAGAAGGTATCAAGGGTGCTTCGTGTTCTTGTTTTTGATTGGTTCTATGTCGGGTGCCATT 495
          || | ||| | |||| ||| |||| || | || | ||||||| |||||| ||| | |||
Db   272 CTCATGGGAGCGCAAGGTTCCGGCAGCATTGTCATGGTAGTTTGGCATGTCGTTGTTCATCGGGTATC 331

Qy   496 GTCAACGANTCCTCAAAAACCAAGTTGTCTACACACATGTCCAAGGCTGGTGTTACCCATTG 555
          ||||||| | || | || || ||| |||| || ||| ||||||||| | || | || |
Db   332 GTCAACGTGCCGCGAGCCCCAGCCGTCCTTACAACAACGGCCGCCGCCCAAGGCCGCCGTTCGTCAACTT 391

Qy   556 GCTAAGACTTTGGCTTGTTGTGAATGGGCTAAGTACAACATCAGAGTTAATTCTTTAAACCCA 615
          || | | | |||| |||||||||||||| | | || | || | | | | | | ||
Db   392 GCCGCGTCCTTCGCCGTGAATGGGCCGCGAGCACCGGGTGAACTGCAACAGCCCT 451

Qy   616 GGTTACATCTACGGTCCTTTTGACCAAGAATGTTATCAATGGTAACGAAGAAATGTACAAC 675
          || |||| | | | | ||||||| || | || | | ||| | | | || | |
Db   452 GGATACAACGTTACTGCCCTGACCGCCAAGAGATTTGGATGAGAACACCCGAAATTGCGGGAC 511

Qy   676 AGATGGGATCTCTGGTATCCCACAACAAAGAAATGTCCGAACACAAAGGAATACAATTGGTGCT 735
          ||||||| | | | | |||||||| ||| ||| ||| | | | | | |||| | |||| |
Db   512 AAGTGGGANTCTGCTCATCCCACCGGCAAGAAGATGGGTACTCCCGAGGACCTGAATGGGTCCC 571

Qy   736 GTTTTGTACTTGCTTTCTGA 755
          ||| |||| | || |||||
Db   572 GTTACCTTCCTGCTCAGTGA 591

RESULT 13
CO141255
LOCUS   CO141255   447 bp   mRNA   linear   EST 17-JUN-2004
DEFINITION   EST835926 Aspergillus flavus Normalized cDNA Expression Library
               Aspergillus flavus cDNA clone NAFEA37 5' end, mRNA sequence.
ACCESSION   CO141255
VERSION   CO141255.1   GI:48890246
KEYWORDS   EST.
SOURCE   Aspergillus flavus
ORGANISM   Aspergillus flavus
               Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
               Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE   1 (bases 1 to 447)
AUTHORS   Yu,J., Whitelaw,C.A., Nierman,W.C., Bhatnagar,D. and Cleveland,T.E.
TITLE   Aspergillus flavus expressed sequence tags for identification of

RESULT 14

```
JOURNAL     FEMS Microbiol. Lett. (2004) In press
COMMENT     genes with putative roles in aflatoxin contamination of crops
            Contact: Yu J
            Food and Feed Safety Research Unit
            USDA/ARS, Southern Regional Research Center
            1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA
            Tel: 504 286 4405
            Fax: 504 286 4419
            Email: jiuyu@srrc.ars.usda.gov
            Contact Dr. Yu at USDA/ARS SRRC (jiuyu@srrc.ars.usda.gov) for clone
            information
            PCR Primers
            FORWARD: M13F
            BACKWARD: M13R
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..447
                     /organism="Aspergillus flavus"
                     /strain="NRRL 3357"
                     /mol_type="mRNA"
                     /db_xref="taxon:5059"
                     /clone="NAFEA37"
                     /sex="asexual mycelia"
                     /cell_type="mycelia"
                     /dev_stage="developmental stages from 18 to 96 hours"
                     /lab_host="E. coli DH10B T1 resistant cells"
                     /clone_lib="Aspergillus flavus Normalized cDNA Expression
                     library"
                     /note="Vector: pBlueScript (SK+) (Stratagene), antibiotic
                     selection marker: Carbenicillin; Site 1: NotI, at the 5
                     prime end; Site 2: EcoRI, at the 3 prime end; This
                     normalized cDNA expression library was constructed using a
                     mixture of mycelial cells grown under eight different
                     medium conditions and harvested at 5 time points (18, 24,
                     48, 72, 96 hours). The poly-A sequence was trimmed off
                     before ligating to vector."
ORIGIN

Query Match     11.9%;  Score 97.2;  DB 7;  Length 447;
Best Local Similarity   58.7%;   Pred. No. 1.4e-16;
Matches 168;  Conservative   0;  Mismatches 118;  Indels   0;  Gaps   0;

Qy    464 TTGTTTTGAATTGGTTCTATGTCTGGTGCCATTGTCAACGGATCCTCAAAACCAAGTTGTGTCT 523
              ||| ||||| |||| ||| | | ||||| | ||||| ||| |  | | | || |||| |
Db    162 TTGTCATGAATTGGTTAGCATTGTCTGGTGTCGTCCGCACGAGGCCCCAGCCTCCTT 221

Qy    524 ACAACATGTCCAAGGCTGGTGTTATCCAATTGGCTAAGGACTTTGGCTTGTGAATGGGCTA 583
              ||||| ||||| |||| ||||| | ||  ||| | | | |  | |||| || ||||||
Db    222 ACAACGCCCCAAGGCCCGCTGTTCGGTGTCAACTCGCCGTCGAATGGGCCG 281

Qy    584 AGTACACAACATCAGAGTTAATTCTTAAACCCAGGTTACATCTACGGTCCTTTGACCAAGA 643
              |||| |||| |||| | || ||| |  ||| || ||| | ||||| ||| ||| ||||
Db    282 GTCACGACATCCGGGTGAACTGCATCAGCCCTACTTACTGCCCTGACCCGCA 341

Qy    644 ATGTTATCAAATGGTAACGAACAGAATGTACAACAGATGGAGTCTCGGTATACCCACAACAAA 703
              | ||| || | |  ||  | || ||| | || | || | |  || | || ||| || |||
Db    342 AGATTTTGGATGAGAGAACCCCGAATTGCCGGGACTGGAGTCTGCCTCATCCCCACCGGCA 401

Qy    704 GAATGTCCGAACCCAAAAGGAATACATTGGTGGTCGCTGTTTTGTACTTCCT 749
              |||| |  |  |||| ||| || | | |||| ||| | ||| | | || ||
Db    402 AGARTGGGTACTCCCGAGGACCTGATGGGTCCCGTTACTTCCTCTGCT 447

RESULT 14
LOCUS       AU249728                 511 bp    mRNA     linear   EST 22-APR-2004
DEFINITION  AU249728 SL Lolium multiflorum cDNA clone SL002A09-5, mRNA
            sequence.
ACCESSION   AU249728
VERSION     AU249728.1  GI:46506997
KEYWORDS    EST.
SOURCE      Lolium multiflorum (Italian ryegrass)
```

RESULT 15

```
ORGANISM    Lolium multiflorum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Poeae; Lolium.
REFERENCE   1 (bases 1 to 511)
  AUTHORS   Ikeda,S.
  TITLE     Lolium multiflorum EST Project
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Seishi Ikeda
            Japan Grassland Farming Forage Seed Association(JFSA)
            Forage Crop Research Institute(FCRI)
            Higashikaba 388-5, Nishinasuno, Tochigi 329-2742, Japan
            Tel: 81-287-37-6757
            Fax: 81-287-37-6755
            Email: sikeda67@ifsass.or.jp
            contact:Tadashi Takanizo (takanizo@affrc.go.jp)
            National Institute of Livestock and Grassland Science, Nishinasuno
            Resistance gene analog.
FEATURES             Location/Qualifiers
     source          1..511
                     /organism="Lolium multiflorum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4521"
                     /clone="SL002A09-5"
                     /tissue_type="Powdery mildew infected leaf"
                     /clone_lib="SL"
ORIGIN

Query Match     11.4%;  Score 93.2;  DB 1;  Length 511;
Best Local Similarity   58.5%;   Pred. No. 2.1e-15;
Matches 200;  Conservative   0;  Mismatches 138;  Indels   4;  Gaps   2;

Qy    469 TTGATTGGTTCTATGTCTGGTGCCATTGTCAACGGATCCTCAAAACCAAGTTGTCTACAAC 528
              || ||||  | ||||| || | ||| |||| | |||  | | || | ||| ||| ||||
Db      2 TTTATTGGAAGCATGTCAGGGACCCGTCGTGAAATGTTCCACAACCACAAGACACCATATAAC 61

Qy    529 ATGTCCCAAGGCTGGTGTTATCCATTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGTAC 588
              | | | |||| | | |||  | | ||||| |||| | ||| | | |||||||| ||  |
Db     62 GCCGGTAAAGGTGCAATTCCGGCACCTGGCATCGAGTGCTACGAGTGGAGTGGGCTCATGCA 121

Qy    589 AACATCGAGTCAATTGTAATTCTTTAAACCCAGGTTACATCTACGGTCCTTTGACCAAGAATGTT 648
              |||||| | | ||||| |  | |  |||| || ||| | || ||| || || || ||||| ||
Db    122 GGAATTCGAGTCAATTGTATTTCGCCTGGAATACATGCTCACTGACTGCAATTGACCAAAAAAAT 181

Qy    649 A-TCAATGGTTAACGAAGAAATTGTACAACAGATGGAGTCTCGGTATCCCACAACAAAGAAT 707
              | | |||| |||| | ||||| || | ||| ||| | |  ||| || ||||| ||| |||
Db    182 ACTCGAAAGAGAAATCCTGATTTAAAGAAGCAAATGGATATTCCGTATTCCACAAGGACAAAAT 241

Qy    708 GTCCGGAACCAAAGAGAAATACATTGGTGGTCGTGTTTTGTACTTGCTGGTCTCTGCTTC 767
              || | | | ||| ||| || || || |  |  | || | ||| | | || |||||
Db    242 GGGTAATCAGAAGAGANTTTGAGTGGGAGCAGTTACTTTCTTGGCTTCTGA---TGCCTCATC 298

Qy    768 ATACACTACTGGGTGCCAGCTTACTCGGTTGANGGGTGGTTTCAC 809
              | |||| |||| |||||| |||||||| | | | ||||| || ||
Db    299 ATACANTGACGGGGCGGGGGGTCGATCTAAGAGGTGGATGGGGGGCTACAC 340

RESULT 15
LOCUS       CN811863                 892 bp    mRNA     linear   EST 01-JUN-2004
DEFINITION  CN811863 Fg09_08p20_A Fg09 AAFC_ECORC_Fusarium_graminearum_simple_substrate
            Gibberella zeae cDNA clone Fg09_08p20, mRNA sequence.
ACCESSION   CN811863
VERSION     CN811863.1  GI:47835874
KEYWORDS    EST.
SOURCE      Gibberella zeae
  ORGANISM  Gibberella zeae
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE   1 (bases 1 to 892)
  AUTHORS   Watson,R.J., Heys,R., Couroux,P., De Moors,A., Harris,L.J.,
            Hattori,J., Lacroix,C., Maotti,M., Ouellet,T., Robert,L.S.,
            Singh,J.A., Sprott,D. and Tinker,N.A.
```

```
TITLE       A cDNA library prepared from Fusarium graminearum grown on a simple
            substrate
JOURNAL     Unpublished (2003)
COMMENT     Contact: Watson, Robert.J.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
            CANADA
            Tel: (613) 759-1655
            Fax: (613) 759-1701
            Email: watsonrj@agr.gc.ca.
FEATURES             Location/Qualifiers
     source          1..892
                     /organism="Gibberella zeae"
                     /mol_type="mRNA"
                     /strain="DAOM 180378"
                     /db_xref="taxon:5518"
                     /clone="Fg09_08p20"
                     /tissue_type="Mycelium"
                     /dev_stage="Asexual"
                     /lab_host="E. coli DH10B"
                     /clone_lib="Fg09_AAFC_ECORC_Fusarium_graminearum_simple_su
                     bstrate"
                     /note="Vector: pBluescript II+; Site_1: EcoRI; Site_2:
                     XhoI; Fusarium graminearum grown on a simple substrate--
                     minimal media supplemented with amino acids."
ORIGIN
Query Match             11.1%;  Score 90.8;  DB 7;  Length 892;
Best Local Similarity   54.5%;  Pred. No. 1.2e-14;
Matches  204;  Conservative   0;  Mismatches 167;  Indels    3;  Gaps    1;

Qy        437 TCAAAGAAAGGTATCAAGGGTGCTTCTTCTGTTGGTTCTTAGTTCTGGTGCCATTG 496
                |||     |     |       |    |  |||| |     ||||||||| ||||||
Db        388 TCAGTGGACGGTCAGGTGCCCGGTAGTATTGTGGTTATTGGAAGCATTGTCTGGTTGCCATTG 329

Qy        497 TCAACGATCCTCAAAACCAAGTTGTCTACAACATGTCCAAGGCTGGGTTTATCCATTTGG 556
                |||||             ||||
Db        328 TCAATGTCCCACAGCCACAAAGCCCCATATAACCGCGGCAAAAGCAGCTTGTTCGACACCTGG 269

Qy        557 CTAAGACTTTGGCTTGTGTGAAATGGGCTAAGTACAAACATCAGAGTTAATTCTTTAAACCCAG 616
                 |    ||
Db        268 CTGCTTCTCTTGCAGTGGAGTGGGCCTCACGCTGGAAATCCGTGTCAAACTGCAATCTCTCCTG 209

Qy        617 GTTACACTACGGTCCTTTGACCAAGAAATGTTATCAATGGGTAACGAAGAATTGTACAAACA 676
                ||||
Db        208 GCTATATGTTGACTGCTTTGACACAGAAGATCCTTCGACGACAACCCTGATTTAGAGAGAA 149

Qy        677 GATGGATCTCTGGTTATCCCACAACAAGAATGTCCGAACCAAAGGAATACATTGGTGGCTG 736
                 |||                  ||         |           || ||
Db        148 CCTGGACATCCCTCATTCCTCAGGGTCGGCATGGGACTGCCTCAAGATTTGATGGGACCCG 89

Qy        737 TTTTGTTACTTGCTTTCTGAAATCGTCGTTCATTACACTACTGGTCTGTGCCAGCTTACTGGGTTG 796
                |||||  |||                  |           |||||
Db         88 TAACCTTTCTGCTTATCAGA---TGCGTCATCTTATATGACTGGGGCAGATGTTCGAGTTG 32

Qy        797 ATGGTTGGTTTCACT 810
                ||||
Db         31 ATGGAGGATACACT 18

Search completed: March  2, 2005, 16:54:05
Job time : 3394 secs
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: March 2, 2005, 15:57:37 ; Search time 594 Seconds
                                           (without alignments)
                                           8140.026 Million cell updates/sec

Title: US-10-720-018-1
Perfect score: 816
Sequence: 1 atgactgactacacattccaac...........atggtggttcacttcttgg 816

Scoring table: IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0
Searched: 9394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications_NA:*
    1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
    2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
    3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
    4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
    5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
    6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
    7: /cgn2_6/ptodata/2/pubpna/PCTUS_NEW_PUB.seq:*
    8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
    9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
   10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
   11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
   12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
   13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
   14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
   15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
   16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
   17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
   18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
   19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
   20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
   21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
   22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	816	100.0	816	18	US-10-720-018-1	Sequence 1, Appli
2	99.4	12.2	3203	18	US-10-653-047-41	Sequence 41, Appl
3	88.6	10.9	637	18	US-10-653-047-4932	Sequence 4932, Ap
4	81.8	10.0	637	17	US-10-369-493-36723	Sequence 36723, A
5	80.6	9.9	735	17	US-10-369-493-26991	Sequence 26991, A
6	74.6	9.1	852	10	US-09-734-237B-71	Sequence 71, Appl
7	74.2	9.1	1146	18	US-10-653-047-4825	Sequence 4825, Ap
8	72.4	8.9	775	18	US-10-425-115-172450	Sequence 172450,
9	67.6	8.3	783	17	US-10-369-493-25810	Sequence 25810, A
10	63.6	7.8	774	17	US-10-369-493-26610	Sequence 26610, A
11	60.6	7.4	636	18	US-10-653-047-3866	Sequence 3866, Ap
12	59.8	7.3	822	15	US-10-314-394-1	Sequence 1, Appli
13	56.4	6.9	825	15	US-10-314-394-3	Sequence 3, Appli
14	56.2	6.9	789	9	US-09-938-842A-948	Sequence 948, App
15	56.2	6.9	789	11	US-09-938-842A-948	Sequence 948, App
16	55.2	6.8	732	17	US-10-369-493-38959	Sequence 38959, A
17	55.2	6.8	732	17	US-10-369-493-38960	Sequence 38960, A
18	55.2	6.8	837	18	US-10-425-115-100705	Sequence 100705,
19	55	6.7	949	9	US-09-770-445-339	Sequence 339, App
20	54.4	6.7	783	9	US-09-938-842A-1025	Sequence 1025, Ap
21	54.4	6.7	783	11	US-09-938-842A-1025	Sequence 1025, Ap
22	53.8	6.6	768	17	US-10-369-493-26631	Sequence 26631, A
23	53.2	6.5	852	18	US-10-639-159-39	Sequence 39, Appl
24	53.2	6.5	852	18	US-10-639-159-53	Sequence 53, Appl
25	53.2	6.5	852	18	US-10-782-258-53	Sequence 53, Appl
26	53.2	6.5	855	18	US-10-639-159-45	Sequence 45, Appl
27	53.2	6.5	855	18	US-10-639-159-47	Sequence 47, Appl
28	53.2	6.5	855	18	US-10-782-258-39	Sequence 39, Appl
29	53.2	6.5	855	18	US-10-782-258-45	Sequence 45, Appl
30	53.2	6.5	855	18	US-10-782-258-47	Sequence 47, Appl
31	52.8	6.5	819	15	US-10-259-165-699	Sequence 699, App
32	52.4	6.4	656	9	US-09-770-149-497	Sequence 497, App
33	52.2	6.4	852	18	US-10-639-159-51	Sequence 51, Appl
34	52.2	6.4	852	18	US-10-782-258-51	Sequence 51, Appl
35	51.6	6.3	852	18	US-10-639-159-55	Sequence 55, Appl
36	51.6	6.3	852	18	US-10-639-159-57	Sequence 57, Appl
37	51.6	6.3	852	18	US-10-782-258-55	Sequence 55, Appl
38	51.6	6.3	852	18	US-10-782-258-57	Sequence 57, Appl
39	51.6	6.3	852	18	US-10-782-258-113	Sequence 113, App
40	51.6	6.3	852	18	US-10-782-258-119	Sequence 119, App
41	51.6	6.3	855	18	US-10-639-159-43	Sequence 43, Appl
42	51.6	6.3	855	18	US-10-639-159-43	Sequence 43, Appl
43	51.4	6.3	849	16	US-10-032-585-6746	Sequence 6746, Ap
44	51.2	6.3	711	15	US-10-259-165-770	Sequence 770, App
45	51.2	6.3	810	9	US-09-938-842A-1690	Sequence 1690, Ap

ALIGNMENTS

RESULT 1
US-10-720-018-1
Sequence 1, Application US/10720018
Publication No. US20040132074A1
GENERAL INFORMATION:
; APPLICANT: Verho, Ritva
; APPLICANT: Richard, Peter
; APPLICANT: Penttila, Merja
; APPLICANT: Verho, Ritva
; TITLE OF INVENTION: New Enzyme for an in vivo and in vitro Utilisation of
; TITLE OF INVENTION: carbohydrates
; FILE REFERENCE: 2530-120
; CURRENT APPLICATION NUMBER: US/10/720,018
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Ambrosiozyma monospora
US-10-720-018-1

  Query Match   100.0%; Score 816; DB 18; Length 816;
  Best Local Similarity   100.0%; Pred. No. 2.4e-230;
  Matches 816; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGACTGACTACACATTCCAACTTTTAGATTCGATGGCCACTTAACCATTGTCACAGGTGCC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGACTGACTACACATTCCAACTTTTAGATTCGATGGCCACTTAACCATTGTCACAGGTGCC 60

Qy   61 TGTGGTGGTTTAGCTGAAGCTTTAATCAAGGGTTTGTTGGCCTACGGTTCTGACATTGCT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGTGGTGGTTTAGCTGAAGCTTTAATCAAGGGTTTGTTGGCCTACGGTTCTGACATTGCT 120

Qy   121 TTGCTTGATATCGACCAAGAAAAGACTGCTGCCAAACACAGCCGGAATACCACAAAATACGCT 180

```
US-10-653-047-41

Qy  121 TTGCTTGATATCGACCAAGAAAAAGAACTGCTGCCAAACAAAGCCGAATACCACAAAATACGGCT  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ACTGAAGAAATTGAAGTTGAAGAAGTTCCAAAAGAAGGGTTCATAATGCCTGTGATATTTCT  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181 ACTGAAGAAATTGAAGTTGAAAGAAGTTCCAAAGATGGGTTCAATATGCCTGTGATATTTCT  240

Db  241 GATTCTGATACCCGTTCACAAGGTGTTTTGCTCAAGGATTTTGGGTAAGTTGCCA  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241 GATTCTGATACCCGTTCACAAGGTGTTTTGCTCAAGTTGCTAAGGAATTTTGGTAAGTTGCCA  300

Db  301 TTGCACTTGGTTAACACAGCTGGTTACTGTGAAAACTTCCCATGTGAAGAAGATTACCCAGCC  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 TTGCACTTGGTTAACACAGCTGGTTACTGTGAAAAACTTCCCATGTGAAGAAGATTACCCAGCC  360

Db  361 AAGAACGCTGAGAAGAATGGTGAAGGTTAACTGTTGTGGGGTTCTTTGTATGTTTCTCAAGCC  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361 AAGAACGCTGAGAGAGATGGTGAAGGTTAACTGTTGTGGGGTTCTTTGTATGTTTCTCAAGCC  420

Db  421 TTTGCTAAGCCACCATTGATCAAAGAAGGTATCAAGGGGTGCTTCTTCGTTTTCTTTGATGGTTCT  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  421 TTTGCTAAGCCATTGATCAAAGAAGAAGGTATCAAGGGTGCTTCTTCGTTTTCTTTGATGGTTCT  480

Db  481 ATGTCTGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTTCTACAACATGTCCAAGGCT  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  481 ATGTCTCTGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTTCTACAACATGTCCAAGGCT  540

Db  541 GGTGTTGATCCCATTTGGGCTAAGACTTTGGCTTGTGAAATGGGTTAACAACATCAGAGTT  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  541 GGTGTTATCCATTGGGCTAAGACTTTGGCTTGTGAAATGGGTTACAACATCAGAGTT  600

Db  601 AATTCTTTAAACCCAGGTTACATCTCACGGTCCTTTTGACCAAGAAGTGTTATCAAGGAAGAAT  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  601 AATTCTTTAAACCCAGGTTACATCTCACTGTTGACCAAGAAGTTGTTATCAACTGGTTAAC  660

Db  661 GAAGAATTGTACAACAGATGGATCTTCTGGTATCCCACAACAAAGAATGTCCGAACCAAAGTACA  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  661 GAAGAATTGTACAACAGATGGATCTTCTGGTATCCCACAACAAAGAATGTCCGAACCAAAAG  720

Db  721 GAATACAATTGGTGGTCGTTTTGTTACTGCTTTCATACACTACTACTGGT  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  721 GAATACAATTGGTGGTCGTTTTTGTTACTGCTTTCATACACTACTACTGGT  780

Db  781 GCCAGCTTACTGGTTGATGGTGGTTTCACTTCTTGG  816
        ||||||||||||||||||||||||||||||||||||
Qy  781 GCCAGCTTACTGGTTGATGGTGGTTTCACTTCTTGG  816
```

; ORGANISM: Fusarium venenatum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(3203)
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-41

Query Match   12.2%; Score 99.4; DB 18; Length 3203;
Best Local Similarity 52.7%; Pred. No. 4.3e-18;
Matches 265; Conservative 0; Mismatches 231; Indels 7; Gaps 2;

```
Qy  308 TGGTTAACACAGCTGGTTACTGTGAAAACTTCCCATGTGAAGAAGATTACCCAGCCAAGAACG  367
        ||||||||||||||||||||||||| |  ||  || | |  |  |||  |    |||||  ||
Db 2614 TAGTCACGTCGGCTGCCTTCACGGAGCAATTCGAGCAATCAACTATCCCATCGATCGTA  2673

Qy  368 CTGAGAAGAATGGTGAAGGTTAACTTGTTGTGGGGTTCTTTGTATGTTTCTCAAGCCTTTGCTA  427
        | ||| | | ||  | |  |  |  |||     || | | ||  | ||| | ||| ||| | |
Db 2674 TGCGCGAAGTTGTGGGGCGTGTTAAAGTTGACGGGTACTTACCTCTTGCAGTTGCAGTTGCCA  2733

Qy  428 AGCCATTGATCAAAGAAGAAGGTATCAAGGGGTGCTTCTTCGTTTTCTTTGATGGTTCT  487
        ||| |||| | ||  |  |  || ||| | || ||||||  |||| | | ||  |||| 
Db 2734 A---ACATCTCATCGGAACGCCAGGTACCTGGTAGTATCCGTGGTTATTGGAAAGCATGTCTG  2790

Qy  488 GTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTTCTACAACATGTCCAAGGCTGGTGTTA  547
        | |||| | |||  ||  | ||  | || || | | |  | || | |  || ||  | | 
Db 2791 GTGCTATTGTCAATGTTCCACAGCCACATATAACGCGGCCAAGGGCAGCGTGTTC  2850

Qy  548 TCCATTTGGCTAAGACTTTGGCTTGTGAAATGGGTTAACAACATCAGAGTTAATTCTT  607
        ||||| | || ||  | ||| || |||| ||||| ||| || | ||| ||| | ||
Db 2851 GTCACCTGGCTGCTTCCCTCGCAGTAGAGTGGGGCTCACGCTGGAATCGAAGTTAAT  2910

Qy  608 TAAACCCAGGTTACATCTCACGGTCCTTTTGACCAAGAAGTGTTATCAAGGAAGAAGAAT  667
        |||   || || || |   | |  | | | | |   |   | ||  || ||  || ||
Db 2911 TCTCTCCTGCGTACAGTTGACTGCTCTAATTCCTCAACGGACAAGAATTCGAACAGATA  2970

Qy  668 TGTACAACAGATGGATCTTCTGGTATCCCACAACAAAGAATGTCCGAACCAAAGAATACA  727
        |||  |  ||||| |  | | | | |  |  ||  |  |  | |||| || || | | |
Db 2971 TCGAGAGGACCTGACATCTCTTTAATTCCTCAAGGACGTATGGGACTACCTCAAGATCTGA  3030

Qy  728 TTGGTGCTGTTTTGTTACTTGCTTTTCGAATCTGCTGCTTCATACACTACTACTGGTGCCAGCT  787
        |  | | |||| | | |  |  ||| |  |    |   |  | |   | |  | | ||||| |
Db 3031 TGGGTCCTGT----GACCTTTTGTTGCAGATGGGTCTTCTCTTACAGTGCCAGGGGCCAGATC  3086

Qy  788 TACTGGTTGATGGTGGTTTTCACT  810
        || |  ||  | |  |  ||  ||
Db 3087 TTAGAGTAGAGTGGAGGAATATACT  3109
```

; ORGANISM: Aspergillus oryzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(687)
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-4932

Query Match   10.9%; Score 88.6; DB 18; Length 687;
Best Local Similarity 49.8%; Pred. No. 2.8e-15;
Matches 279; Conservative 0; Mismatches 275; Indels 6; Gaps 2;

```
Qy  196 TTGAAAGAAGTTCAAAAGATGGGTTCATTATGCCTGTGATATTTCTGATTCTGATACCGGTT 255
            ||||||| ||||| |  | || |||| | | | || ||| || | |||||| || ||
Db  114 TTGGAACAAATGCCCAAGGTCACCGGCCCCAGTACCGCTGATGTTTCCGACCCTAACTCCGTC 173

Qy  256 CACAAGGTGTTGCTCAGTTGCTTGCTAAGGATTTTGGTAAGTTGCCATTTGCACTTGGTTAAC 315
            | |||||| | | | |||| |||| | |||| ||| |||| | ||| ||  |||||| ||
Db  174 AACGGATGCCCTCCTCCGATATTATCTCCAAGCACCGGCCAAGAT---CGACAAACCTGGTCACC 230

Qy  316 ACAGCTGGTTACTGTGAAAAACTTCCCATGTGAAGGAACGCTGAGAG 375
            || ||| |||||  |||| ||| ||| || | |||| ||  |||||
Db  231 TCCGCCGGATTCACGGGAAAACTTCGATGCCCATCTTCCTACCCTCACGACGTCTGCAAAAG 290

Qy  376 ATGGTGAAGGTTAACTTGTTGGGGTTCTTTTGTTATGTTTCTCAAGGCCTTTGCTAAGCCATTG 435
            ||||| ||  || || ||||| | ||  |||| | | | ||| | ||| ||  ||||| |||
Db  291 CTTTGGGGCGGTTAATGTCGATGGAACATATCTTTTCGCCACCGGTGTCGGCCAAGC---AC 347

Qy  436 ATCAAAGAAGGTATCAAGGGGTGCTTCTGTGTTGTTCTATGTGTCTCGGTGCCATT 495
            ||| || |||| | || |||||| | || | ||| | ||||| |||||| ||||
Db  348 CTCATGGACGGCACAAGGTTCCGGGGCAGCAATGGTACCATGTCTCGGTGCTATC 407

Qy  496 GTCAACGATCCTCAAAACCAAGTTGTCTACAACATGTCCAAAGGCTGGTGTTATCCATTTG 555
            |||| | ||||  | | |||| ||| | |||| ||| |||||  || |||| | |||||
Db  408 GACAACGTTGCCCAGCCCAGGCTTCCTTACACAACGCCGCCGTGTTTCGGGGTTCCGGGTG 467

Qy  556 GCTAAGACTTTGGCTTGTTGAATGGGCTAAGTACAACATCAGAGTTAATTCTTTAAACCCA 615
            |||| |||||| ||| |||||| ||| | ||||||| ||||| ||||| ||| ||||||
Db  468 GCCCGCTACTTCGGCGGCGGAATGGGCCGGTCACGACATCCGGTGAGCTGCATGAGCCCT 527

Qy  616 GGTTACAICTCACCGGTCCTTTGACCAAGAATGTTTATCAACGAAGAATGTGTACAAC 675
            ||||  | | | || || ||| |||||| | || | || ||  | || ||  | ||
Db  528 CGAACACGCTTACTGCCCTGACCCGCAAGATTTTGATGAGGAACCCCGATTGGGGAC 587

Qy  676 AGATGGATCTCTGGTATCCCACACAAACAGAATGTCCGGACCAAAAGGAATACATTGGTGGTGCT 735
            |||| |||| ||| | ||||| |  |  |||| || | |||| ||  || | |||||| |||
Db  588 AAGTGGATCTCGCTCATCCCCACCCGGGGTACTCCCGAGGACCTAGGGGTCCC 647

Qy  736 GTTTTGTACTTGCTTTCTGA 755
            |||||||||| ||  |||
Db  648 GTTACCTTTCTGCTCAGNGA 667
```

RESULT 4
US-10-369-493-36723
Sequence 36723, Application US/10369493
; Publication No. US20030236750A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 36723
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-10-369-493-36723

Query Match   10.0%; Score 81.8; DB 17; Length 637;
Best Local Similarity 55.0%; Pred. No. 2.7e-13;
Matches 183; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

```
Qy  304 CACTTGGTTAACACAGCTGGTTACTGTGAAAAACTTCCCATGTGAAGAATTACCCAGCCAAG 363
            ||||| || | || |||||| |||| | |||| || ||| || |  |||| | ||||||
Db   70 CATCTCCGTCACGTCTGCGCGGGATTTACCGGAAAACTTCGACGGCCATCAGCGTACCGAC 129

Qy  364 AACGCTGAGGAGAGATGGTGAAGGTTAACTTGTTGGGGTTCTTTGTATGTTTCTCAAGCCTTT 423
            || ||||| |   || ||| | |||| ||||  | ||| ||  | | | ||| | ||  |||
Db  130 CGTATGCAGAGAGCTCTGGGGGGTGAACGTTGGACGGCACATATCTGTTTGGCTACGGCGTG 189

Qy  424 GCTAAGCCAATTGATCAAAGAAGGTATCAACAAGGTGCTTCTGTGTTTTGATTGGTTCTATG 483
            |||| || ||||| |  ||  |||| |  | ||| || || | | || || || | ||||
Db  190 GCAAAGCACTTGAT---GGAGCGCAAGAGCGCCCCGGAGTAATTGGTGGATGGTAGCATG 246

Qy  484 TCTGGTGCCATTGTCAACGGATCCTCAAAAACCAAGTTGTTCTACAACATGTCCAAGGCTGGT 543
            |||| || ||||| |||| |||| || | | ||| ||| | ||||| ||| |||| ||||
Db  247 TCTGGTGCCATCGGTCAAAGTTACCTCAAACCACAACGGCGCCCCTTACACACGCGCCAAGGCAGCC 306

Qy  544 GTTATCCATTTGGGCTAAGACTTTGGCTTGTTGTGAAATGGGCTAAGTACAACATCAGAGTTAAT 603
            |||| | ||| ||| | |||| | ||| |||| | ||||| | || |||||| || |||||
Db  307 GTTCGCCACCTTGCTTCGCCGTCGTTGTCCGCGTCGAAATGGGGCAGGACGCATGGTATCAGGTAAAC 366

Qy  604 TCTTTAAACCCAGGTTAACATCTACAGGTCTTTG 636
            ||||  |||| |||| | |  || || |  ||||
Db  367 TGCATTAGCCCTGGATACAATACATCCTGACTGCTCTG 399
```

RESULT 5
US-10-369-493-26991
Sequence 26991, Application US/10369493
; Publication No. US20030236750A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26991
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Neurospora crassa
US-10-369-493-26991

Query Match   9.9%; Score 80.6; DB 17; Length 735;
Best Local Similarity 61.8%; Pred. No. 6.7e-13;
Matches 128; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
Qy  436 ATCAAAGAAGGTATCAAGGGGTGCTTCTGTGTTGTTCTATGTGTCTCGGTGCCATT 495
            |||| | |  || | | ||| ||| |  ||| | | ||| | |||| |||| |||
Db  364 ATGATGGCCGCTGGCGAAACCCGGCGCTCCATTATCCTTGTTGCCTCGANTGCGGCCATC 423

Qy  496 GTCAACGATCCTCAAAACCAAGTTGTCTACAACATGTCCAAGGCTGGTGGTTGTTATCCATTTG 555
            ||||| | | | | | | |||| | | ||||  | || ||| | | || || || || ||
Db  424 GTCAACTACCCTCAGGAGCAGTCGTGCTACAACAAACCGCCTCTAGGCTCGGTCGTCATTCAGCTC 483

Qy  556 GCTAAGACTTTGGCTTGTTGTGAATGGGCTAAGTACAACATCAGAGTAATTCTTTAAACCCA 615
            |||| |  ||| ||| |   |  | |||| |||||||||||||| || | |   ||||||
Db  484 GGCAAGTCTCTTGCTGCCCAGGACGTGGGCCCAAGTACAACACATCCGTGTTAAACTGCAATTCCCC 543

Qy  616 GGTTACAICTCACCGGTCCTTTGACCAAG 642
            ||||| | | | |||| ||| ||||||||
Db  544 GGTTACAACTTACGGTCCTTTGACCAAG 616
```

Db   544 GGTTACACGGATACCGCCCTGAATAAG 570
          |||||||||| | || ||| |||

RESULT 6
US-09-734-237B-71
; Sequence 71, Application US/09734237B
; Publication No. US20030644432A1
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 71
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic gene derived from Candida magnoliae NADPH-dependent car
; OTHER INFORMATION: bonyl reductase, having numerous codons replaced with others enco
; OTHER INFORMATION: ding the same amino acids to reduce the free energy of folding
US-09-734-237B-71

Query Match   9.1%;   Score 74.6; DB 10; Length 852;
Best Local Similarity   56.0%;   Pred. No. 4.3e-11;
Matches 209; Conservative   0; Mismatches 149;   Indels 15;   Gaps 3;

Qy   437 TCAAAGAAGGTATCAAGGGTGGCTTCTGTTGTTTGATTGGTTGGTTGTATGTCTGGTGCCATTG 496
          | || || | ||| | |||||| || | || |||||| | ||| ||||||||| ||||
Db   485 TCGAAAAGGAAGGTTAAGAAAGGCCCTCTGGTTTTCACCGGTTCCTTCTGTATGTCCACATCG 544

Qy   497 TCAACGATCCTCAAAAACCAAGAGTTGTCTACACAACATGTCCAAGGCTGGTGTTATCCATTGG 556
          ||||||| ||| ||| ||| ||||| | ||||||| ||||||| |||||| || ||||||||
Db   545 TTAACGTACCGCAGTTTCAGGCTACCTTCACACACCTCTAAAGTGGTTGTTCGTCACTTCG 604

Qy   557 CTAAGACTTTGGGCTTGTTGAATGGGCTAAGTAGTACAACATCAGAGTTAATTCTTTAAACCCAG 616
          || | ||||||||| || || ||| ||||| || ||| | |||| |||| |||| ||||||
Db   605 CTAAAATCTCTGGCTGTAGAATTCGGCTCCGT---TCGCTCGTGTTAACTCGTTTCTCCGG 661

Qy   617 GTTACATCTACGGTCCTTTGACCAAGAATGTTATCAATGGTAACGAAGAATTGTTACAACA 676
          || ||| || ||| | |||| || ||| ||| ||| |||||| |||| | ||| ||||
Db   662 GCTACATCAAC--------ACCGAAATCTCTGACTTTTGTACCGCAGGAAAACTCAGAACA 712

Qy   677 GAATGGAATCTCTGGTATCCCACACAACAAGAAGAATCTCCGAACCAAAAGGAATACATTGGGTCTG 736
          |||| ||| || ||| | ||||||| || ||| | | | ||| | ||| || || ||||
Db   713 AATGGTGGTCTCTGGTACCGCTGGGCCCGTGGAAACTGCTGAACTGCTTGGTTGGTGCTT 772

Qy   737 TTTTGTACTTGCTTTCTGAATCTGCTTCATACACTACTGGTGTCCAGCTTACTGGTTG 796
          |||| | |||||| ||| | | |||| | || ||| || || ||| ||||| |
Db   773 ACCTGTTCTCTGGCTTCTGA---CGCTCGGTTCTTACGGCACTGACATCATCGTTG 829

Qy   797 ATGGTGGTTTCAC 809
          ||| |||| ||||
Db   830 ACGGTGGTTTACAC 842

RESULT 7
US-10-653-047-4825
; Sequence 4825, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R.Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen

; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849,200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4825
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
US-10-653-047-4825

Query Match   9.1%;   Score 74.2; DB 18; Length 1146;
Best Local Similarity   50.6%;   Pred. No. 6.7e-11;
Matches 206; Conservative   0; Mismatches 198;   Indels 3;   Gaps 1;

Qy   349 GATTACCAGCCAAGAAACCCTGAGAAGATGGTGAAGGTTAACTTGTTGGGTTCTTTGTAT 408
          |||||||| | | | ||| ||||| |||| |||| ||| | |||| |||| ||| ||
Db   684 CTAATTGCCGAGCATGAGTGGTACTAATTGCCAATCGGGGTGTTCATCGCCCCTGCTTACGCAAT 743

Qy   409 GTTTCTCAAGCCCTTTGCTAAGCCATTGTCAACGCATGTCAAGAGAGGTATCAAGGGTGTTGCTTCTGTTCGTGTTT 468
          | | | || | ||| | ||| |||||| ||||||| |||||| || | |||| | || | | || || ||
Db   567 GAGTATACGGCCCCAGGACGCCCAACACGACGATGTTCGAAGTCAACGTCACTGGTGTGTTCATG 626

Qy   469 TTGATTGGTTCTATGTCTCGGGTGGCCATTGTCAAGGATCCTCAAAAACCAAGTTGTCTACAAC 528
          || | || | | || | ||||| ||||| |||| | | |||||||||| | ||||||||
Db   804 AACATTCGATCAACACCATCTCGCCCGGCTACATTGTTACTGGTCAAGGTGGTGTTTACTTCAC 863

Qy   529 ATGTCCAAGGCTGCTGTGTTTATCCATTGGTCAAGGGTTGGTTGGTTGTGAGAATGGCTAAGTAC 588
          |||||| || ||| ||| || || || || |||| | | || ||| ||| | | |||| |
Db   744 GCTAGCAACGGCTCGCAGTGCTTCAACTTGCCCGTAACCTGCACCATGGAGTGGGGCCCGGTAC 803

Qy   589 AACATCAGAGTTAATTCTTTAAAACCCAGGTTACATCACTACGGGTCCTTTGACCAAGAATGTT 648
          || || |||| || | ||| | | | | ||||| | | ||| | ||||| || ||| ||
Db   589 AACATTCGAGTCAACACCATCTCGCCCGGCTACATTGTTACTGGTCAAGGTGGTGTTTACTTCAC 648

Qy   649 ATCAATGGTAACGAAGAATTGTACAACAGATGATCTCTGGTATCCCACACAACAAGAATG 708
          || ||| ||||| |||| | ||| | ||| || | ||| | |||||| | | || ||
Db   864 TTCGTTGAGTTCCCTGAGCGGTCGCGGAGGAATGGCCCAAACATAACATGCTGGGGACGTCTG 923

Qy   709 TCCGAACCAAAAGGAATACATTGGGTCTGTTTTGTACTTGCTTTCTGA 755
          | | || |||| ||| | || | || | | ||||| | | || |
Db   924 TCTACCCCTAACGAGTACCGTGGCGCGTCCGCGTCTTCCCTTCTCAGTGA 970

RESULT 8
US-10-425-115-172450
; Sequence 172450, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 172450
; LENGTH: 775
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MR14577_888580.1
US-10-425-115-172450

Query Match   8.9%; Score 72.4; DB 18; Length 775;
Best Local Similarity   56.2%;   Pred. No. 1.8e-10;
Matches 136; Conservative   0; Mismatches 106; Indels   0; Gaps   0;

```
Qy  383 AGGTTAACTTGTTGTGGTTGTTCTTGTTGTTTGTTTTGTTCTCAAGGCATTGATCAAAG 442
Db  465 ACGTCAAGGTTTGGGGACAGTTCGCCTGTGTCTCAAGCCACCTTCTGGACCAAGT 524

Qy  443 AAGGTATCAAGGGGTGCTTCTGTTGTTGTTTTGATTGGTTCTTATGTCTGGTGACACG 502
Db  525 CTGGCTTCAAGCACGGTTCCATCCTCACCTCGGTTCCATGTCTTCCAAATCGTCAAC 584

Qy  503 ATCCTCAAAAACCAAGTTGTCTACAAACATGTCCAAGGCTGGTGTTATTCATTGGCTAAGA 562
Db  585 GTGGCATCCACCAAGTCTTCTACAACATCGTTCCAAGGCTGCCGCCTCGGTGCGTTAAGC 644

Qy  563 CTTTGGCTTGTTGAATGGGGGCTAAGTACAACATCAGAGTTAATTTTCTTTAAACCCCAGGTTACA 622
Db  645 AGGTCGGCTGTTGAGTGGGGCCGACAAGGGGTATTCGAGTCAACGCTTCTCTCCCAGGTTACG 704

Qy  623 TC 624
Db  705 TC 706
```

RESULT 9
US-10-369-493-25810
; Sequence 25810, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 25810
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-25810

Query Match   8.3%; Score 67.6; DB 17; Length 783;
Best Local Similarity   51.2%;   Pred. No. 4.8e-09;
Matches 213; Conservative   0; Mismatches 194; Indels   9; Gaps   2;

```
Qy  171 CAAATACGCTACTGAAGAATTGAAGTTGAAAGAAGTTCCAAAGATAGGGTTCATATGCCTG 230
Db  171 CGAAAAGGCTGCAAAAGGAAATTGCCCAGGCTAATCGTGTCTAATACATGCAAGTG 230

Qy  231 TGATATTTCTGATTCTGATACCGTTCACAAGGTGTTGCTCAAGGTTGCTAAGGATTTTGG 290
Db  231 CGATGCTACTATTCCTAAAGAGGTTAGAACATGCCCTTGCTTGGCTGAGATTCTTTGA 290

Qy  291 TAAGTTGCCATTGCACTTGGTTACACACAGCTCGGTTACTCGTTGAAAACTTCCCATGTGAAGA 350
Db  291 TACTATTGATATCGTTGTGTCCCACAACGGTATTTGCACTGGCAAGTCCGCCATT---GA 347

Qy  351 TTACCCACGCCAAGAACGCTGAGAAGAATGGTGAAGGTAAGTTCGAAGGTTCGGGTT 410
Db  348 CATGACTTACGGAAGAATTTGCTAATGAAAATAAAACGTCAACTTGCTCGGGTGTCTC 407

Qy  411 TTCTCAAGCCTTTGCTAAGCCATTGATCAAAGAAGAGAGTATCAAGGGGTGCTTCTGTTTT 470
Db  408 CGCTCACAATGCTGCGCCCCATCTTCAAAAAGCAAGG------CCATGGCAGCCTTGTAGC 461
```

RESULT 10
US-10-369-493-26610
; Sequence 26610, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26610
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Thermotoga maritima
US-10-369-493-26610

Query Match   7.8%; Score 63.6; DB 17; Length 774;
Best Local Similarity   61.4%;   Pred. No. 7.3e-08;
Matches 102; Conservative   0; Mismatches 64; Indels   0; Gaps   0;

```
Qy  471 GATTGGTTCTTATGTCTGGTGGCCATTGTCAACGATCCTCAAAAACCAAGTTGTCTACAACAT 530
Db  462 TACTGCTTCCAATGTCTGGTGGTGTTGTGGTGTTGTTAACGTTCTCAACAACAATGTGCTTACAACAC 521

Qy  531 GTCCAAGGCTGGTGTTATTCATTGGCTAAGACTTTGGCTTGTTGAATTGGGGCTAAGT 586
Db  522 TTCCAAGGCCGGTGTCATTCAACTCATCAAGAGTTTGGGAGTCGAAGATGGGCGCAAAT 577
```

RESULT 11
US-10-653-047-3866
; Sequence 3866, Application US/10653047
; Publication No. US20040029367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods for Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSEQ for Windows Version 4.0

```
Qy  472 ATTGGTTCTTATGTCTGGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTCTACAACATG 531
Db  436 ATCGGTTCCATGTCCGACACATCGTGAACAAACCTCAGAGACGACAGCAGCTTACAACGCT 495

Qy  532 TCCAAGGCTCGGTGTTATCCATTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGTACAAC 591
Db  496 TCGAAAGCGGGGTGTGATCCATCTCACCAGATCTCTCGGCGCCCGGAGTGGGGCCCGTACGGA 555

Qy  592 ATCAGAGTTAATTCTTTAAACCCAGGTTACATCTACGGTTCCTTTGA 637
Db  556 ATCAGGGTGAACAGCATAAGCCCCGGATAACATCAGAACACAGCTCTCA 601
```

```
; SEQ ID NO 3866
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(636)
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-3866

  Query Match      7.4%;  Score 60.6;  DB 18;  Length 636;
  Best Local Similarity  51.3%;  Conservative    0;
  Matches 194;  Conservative    0;  Mismatches 175;  Indels   9;  Gaps   2;

Qy    211 AAGATGGGGTTCATTATGCCTGTGATATTTCTGATTCTGATACCGTTCACGAAGGTGTTTGCT 270
          ||||| ||||| | | |||||||| ||| |||||| |||| ||| |||| ||| |||  |
Db    200 AAGTCAGCTGCCTATAAAAGCTGATGTCCGGAAATCAGGAGGGATAATTGAGAGGGTAATCCAG 259

Qy    271 CAAGTTGCTAAGGAGATTTTGGTAAGTTGCCATTGCACTGGTTAACACAGCTGGTTACTGT 330
          ||||  || |||| |||| ||  | ||| || ||| ||||| ||||||| ||| ||| |||
Db    260 CAGATTGCTTCAGACTTTGGCAAGCTTGATATTATCGTCGTGAAC---TCGGGGGCTCACT 316

Qy    331 GAAAACTTCCCATGTGAAGAATTACCCAGCCAAGAAGCGTGAGAAGAATGGTGAAGGTTAAC 390
          ||| | ||  | |||| || | ||| || | | |||| ||||||||| |||||| || |||
Db    317 TCCAATATCGCGCGAGAAGAACTACACACCGACAAATGCGTGACACATCATGAGAGGTCAAT 376

Qy    391 TTGTTGGGGTTCTTTGTATGTTTCTCAAGCCTTGCTAAGCCATTGATCAAAGAGAGGTATC 450
          ||| | ||||| |||| | |||| |||| || ||| || ||| | ||| | | ||| |
Db    377 CTAGATGGGCGCAATTTATAGCGCGGCAAGGGCTTCTCCGGATCTTCACAAGGACACAGG---- 432

Qy    451 AAGGGTGCGTTCTGTTGTTTGTTTGATTGGTTTCTATGTCTGTGGCCATTGTCAACGATCCTCAA 510
          | |||| || ||||| ||||| | ||| || |||||| | | || |||| | | | || |
Db    433 --ACATGGAAAATGTTAATCTTCACAGCCCTCGTGTCTCAGTGCAACATTGGTGAAGTGCCCTCAG 490

Qy    511 AACCAAGTTGTCTACAACACATGTCCAAGGCTGGTGTGTTATCCATTTGGCTAAGACTTTGGCT 570
          ||| ||||| | | || || ||| | ||||||| |||| | || ||| | || |||||||| |||
Db    491 AAGGACAGTGCATACAATCCCTCCAAAGCTGGGTCGTGCTGCAAAATGGGCAAAAATGTCTGTCT 550

Qy    571 TGTGAATGGGCTAAGTAC 588
          ||||| |||||| ||
Db    551 GTTGAATGGGTTGANTTC 568

Qy    813 T 813
          |
Db    807 T 807
```

```
RESULT 12
US-10-314-394-1
; Sequence 1, Application US/10314394
; Publication No. US2003014370A1
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Hiroaki
; APPLICANT: Ueda, Momoko
; APPLICANT: Pan, Ritsuzui
; APPLICANT: Hamatani, Takeshi
; TITLE OF INVENTION: METHODS FOR PRODUCING OPTICALLY ACTIVE ALCOHOLS
; FILE REFERENCE: SHZ-006
; CURRENT APPLICATION NUMBER: US/10/314,394
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/385,434
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: JP 2002-152955
; PRIOR FILING DATE: 2002-05-27
; PRIOR APPLICATION NUMBER: JP 2001-375041
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 1
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Datura stramonium
US-10-314-394-1

  Query Match      7.3%;  Score 59.8;  DB 15;  Length 822;
  Best Local Similarity  52.5%;  Pred. No. 1e-06;
  Matches 158;  Conservative    0;  Mismatches 137;  Indels   6;  Gaps   1;
```

```
Qy    519 TGTCTACAACATGTCCAAGGCTGGTGTGTTATCCATTTGGGCTAAGACTTTGGCTTGTGAATG 578
          |||||  |||| ||| | |||| |||||| ||  | |||| || |||||||||| || || |||
Db    507 TCTTTACTCAGCTTCCAAAGGTGCAATAAATCAAATGACAAAGAGTTTGGCTTGTGAATG 566

Qy    579 GGCTAAGTACAACAATCAGAGTTAATTCTTTAAACCCAGGTTACATCTACGGTCCTTGAC 638
          |||||| ||||| | |  || | |||  |  |||| ||| |||| | | || | || |
Db    567 GGCTAAAGACAACAATCGGGTCAATTCAGTTGCTTCCCGGGAGTCATTTTAACCCACTGGT 626

Qy    639 CAAGAAATGTTATC------AATGGTAACGAAGAATTGTACAACAGATGGATCTCTGGTTAT 692
          |||| ||| |  |        ||| |||| |||| | |||| | ||| || | ||||| ||
Db    627 TGAAACTGCAATTAAGAAAAATCCTCATCAAAAGAGAAATAGACAATTTATTGTTGTCAA 686

Qy    693 CCCACAACAAAAGAAATGTCCGAACAAACAAAAGGAATACAATTGGTTGCGTTGTTTTGTTTTC 752
          |||| ||||| | || ||| || ||| | ||| || |||| || | || ||| | | || |
Db    687 GACTCCTATGGGCCCGGGAAAGCCCCAAGAGAGTTTCTGCACTAATAGCTTTTCTTTG 746

Qy    753 TGAATCTGCTGCTTCATACACTACTGGTGCCAGCTTACTGGTTGANGGTGGTTTCACTTC 812
          ||||||||| | ||||  |  |  | ||||| | || | | | || || | || ||||
Db    747 CTTCCCTGCTTCATATATTACGGGCCAGATCGATCATATGGGCTGAACGGTGGATCACAGC 806

Qy    813 T 813
          |
Db    807 T 807
```

```
RESULT 13
US-10-314-394-3
; Sequence 3, Application US/10314394
; Publication No. US2003014370A1
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Hiroaki
; APPLICANT: Ueda, Momoko
; APPLICANT: Pan, Ritsuzui
; APPLICANT: Hamatani, Takeshi
; TITLE OF INVENTION: METHODS FOR PRODUCING OPTICALLY ACTIVE ALCOHOLS
; FILE REFERENCE: SHZ-006
; CURRENT APPLICATION NUMBER: US/10/314,394
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/385,434
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: JP 2002-152955
; PRIOR FILING DATE: 2002-05-27
; PRIOR APPLICATION NUMBER: JP 2001-375041
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 3
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Hyoscyamus niger
US-10-314-394-3

  Query Match      6.9%;  Score 56.4;  DB 15;  Length 825;
  Best Local Similarity  52.4%;  Pred. No. 1e-05;
  Matches 151;  Conservative    0;  Mismatches 131;  Indels   6;  Gaps   1;
```

```
Qy    532 TCCAAGGCTGGTGTGTTATCCATTTGGGCTAAGACTTTGGCTTGTGAAATGGGCTAAGTACAAC 591
          |||| || | | || | | |||| || || ||| || ||||| || || ||| |||| ||
Db    523 TCCAAAGCTGCAATAAATCAAATAACGAAGAACTTGGCCATGTGAAATGGGCCAAGGACAAC 582

Qy    592 ATCAGAGTTAATTCTTTAAACCCAGGTTACATCTACGGTCCTTGACCAAGAATGTTATC 651
          || |||||| ||| | |  | ||| ||||| | | | || || || ||||| ||| | |
Db    583 ATTCGGGTCAATTCAGTTGCTCCAGGAGTCATTTTAACCCACTTTAAACCTGCAATT 642

Qy    652 ------AATGGTAACGAAGAATTGTACAACAGATGGATCTCTGGTTATCCCACAACAAAGA 705
                 ||| ||||| ||||  |  | ||| |  ||| || | ||||| || || | | |
Db    643 AAGAAAAATCCTCATCAAAAGAGAAATAGACAATTTATTGTTGTCAAGACTCCAATGGGC 702

Qy    706 ATGTCCGAACCAAAGGAATACAATTGGTTGCGTTGTTTTGTTTTCTGAATCTGCTGCT 765
          || || ||| |||| |  |  | || | || ||| | | || || ||||||| || ||
Db    703 CGGGCTGGAAAGCCCAATGAAGTGTCTGCACTAATAGCCTTTCTTTTGCTTCGCTGCT 762
```

```
Qy  766 TCATACACTACTGCTGCCAGCTTACTGTGGTTGGTTGATGGTGGTTTTCACTTCT  813
            ||||||  ||||||||||   |||||    ||||| |||||||   ||||||
Db  763 TCTTATATTACTGGCCAAATTATATGGGCGTGATGGTGGATTCACAGCT       810
```

; US-09-938-842A-948
; Sequence 948, Application US/0993842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;   APPLICANT: Harper, Jeff
;   APPLICANT: Kreps, Joel
;   APPLICANT: Wang, Xun
;   APPLICANT: Zhu, Tong
;   TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;   FILE REFERENCE: SCRIP1300-3
;   CURRENT APPLICATION NUMBER: US/09/938,842A
;   CURRENT FILING DATE: 2001-08-24
;   PRIOR APPLICATION NUMBER: US 60/227,866
;   PRIOR FILING DATE: 2000-08-24
;   PRIOR APPLICATION NUMBER: US 60/264,647
;   PRIOR FILING DATE: 2001-01-16
;   PRIOR APPLICATION NUMBER: US 60/300,111
;   PRIOR FILING DATE: 2001-06-22
;   NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 948
;   LENGTH: 789
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-948

Query Match   6.9%; Score 56.2; DB 9; Length 789;
Best Local Similarity   55.2%; Pred. No. 1.1e-05;
Matches 153; Conservative   0; Mismatches 118; Indels   6; Gaps   2;

```
Qy  533 CCAAGGCTCGGTGTTATCCATTTGGCTAAGACTTTTGGCTTGTTGAATGGGCTAAGTACAACA  592
            ||||||| ||||  ||||||    ||||| ||| ||||| |||| ||||||| |||||
Db  485 CAAAAGGAGCTTTGATTCAGCTAGCTAAAAAATTTGGCATGTGAATGGGCAAAAGACGGCA     544

Qy  593 TCAGAGTTAATTCTTTAAACCCAGGTTACATCTACGGGTCCTTTGACCAAGAATGTTATCA     652
            |||||||  ||||  ||||||||||| |||||   ||||||||||||||||||||||||
Db  545 TAAGAGCCAACGCTGTTGCGCCTAATGTCATCAATAACTCCTCGTCTCAATCTTATCTTG     604

Qy  653 ATGGTAACGAAGAATTGTACAACAGATGGAATCTCTGGTATCCCACAACAAAGAAATGTCCG   712
            ||  ||||||   ||| ||| ||||||||  ||||||||  | |||||||||| ||||
Db  605 AGG---ACGTCAGTTTCAAGAAGGCCATTGTTGAGTAGGACTCCACTGGTCGTGTTGGGAG   661

Qy  713 AACCAAAGAGAATACATTGGTGCTGTTTTTGTTACTTGCTTTCTGAATCTGTCGTGCTTCATACA  772
            |||||||||||| ||||| || | ||||||||||||| ||  ||||| |||| |||||
Db  662 AGCCAAAATGAAG---TTGCATCACTAGTGGCCTTCTTGTGTCTCACTGCAGCGCTTCTTATA   718

Qy  773 CTACTGGTGCCAGCTTACTGGTTGATGGTGGTTTCAC  809
            |||||||  | |||||| || |||| ||| ||||
Db  719 TTACTGGTTCAGACTATTTGTTGTTGATGGAGGTCTCAC   755
```

; RESULT 15
; US-09-938-842A-948
; Sequence 948, Application US/0993842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
;   APPLICANT: Harper, Jeff
;   APPLICANT: Kreps, Joel
;   APPLICANT: Wang, Xun
;   APPLICANT: Zhu, Tong
;   TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;   TITLE OF INVENTION: SAME, AND METHODS OF USE
;   FILE REFERENCE: SCRIP1300-3
;   CURRENT APPLICATION NUMBER: US/09/938,842A
;   CURRENT FILING DATE: 2001-08-24
;   PRIOR APPLICATION NUMBER: US 60/227,866
;   PRIOR FILING DATE: 2000-08-24
;   PRIOR APPLICATION NUMBER: US 60/264,647
;   PRIOR FILING DATE: 2001-01-16
;   PRIOR APPLICATION NUMBER: US 60/300,111
;   PRIOR FILING DATE: 2001-06-22
;   NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 948
;   LENGTH: 789
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-948

Query Match   6.9%; Score 56.2; DB 11; Length 789;
Best Local Similarity   55.2%; Pred. No. 1.1e-05;
Matches 153; Conservative   0; Mismatches 118; Indels   6; Gaps   2;

```
Qy  533 CCAAGGCTCGGTGTTATCCATTTGGCTAAGACTTTTGGCTTGTTGAATGGGCTAAGTACAACA  592
            ||||||| ||||  ||||||    ||||| ||| ||||| |||| ||||||| |||||
Db  485 CAAAAGGAGCTTTGATTCAGCTAGCTAAAAAATTTGGCATGTGAATGGGCAAAAGACGGCA     544

Qy  593 TCAGAGTTAATTCTTTAAACCCAGGTTACATCTACGGGTCCTTTGACCAAGAATGTTATCA     652
            |||||||  ||||  ||||||||||| |||||   ||||||||||||||||||||||||
Db  545 TAAGAGCCAACGCTGTTGCGCCTAATGTCATCAATAACTCCTCGTCTCAATCTTATCTTG     604

Qy  653 ATGGTAACGAAGAATTGTACAACAGATGGAATCTCTGGTATCCCACAACAAAGAAATGTCCG   712
            ||  ||||||   ||| ||| ||||||||  ||||||||  | |||||||||| ||||
Db  605 AGG---ACGTCAGTTTCAAGAAGGCCATTGTTGAGTAGGACTCCACTGGTCGTGTTGGGAG   661

Qy  713 AACCAAAGAGAATACATTGGTGCTGTTTTTGTTACTTGCTTTCTGAATCTGTCGTGCTTCATACA  772
            |||||||||||| ||||| || | ||||||||||||| ||  ||||| |||| |||||
Db  662 AGCCAAAATGAAG---TTGCATCACTAGTGGCCTTCTTGTGTCTCACTGCAGCGCTTCTTATA   718

Qy  773 CTACTGGTGCCAGCTTACTGGTTGATGGTGGTTTCAC  809
            |||||||  | |||||| || |||| ||| ||||
Db  719 TTACTGGTTCAGACTATTTGTTGTTGATGGAGGTCTCAC   755
```

Search completed: March 2, 2005, 18:17:37
Job time : 597 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: March 2, 2005, 11:05:37 ; Search time 185 Seconds
                 (without alignments)
                 7217.307 Million cell updates/sec

Title: US-10-720-018-1
Perfect score: 816
Sequence: 1 atgactgactacattcccaac.........atggtggttcaacttcctgg 816
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 1202784 seqs, 81818359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq ldength: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   !Issued_Patents_NA:*
                 1:/cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 1:/cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3:/cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4:/cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5:/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 6:/cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Match Length	DB	ID	Description
	1	238.2	29.2	696	2	US-08-336-198C-6	Sequence 6, Appli
	2	238.2	29.2	696	2	US-09-184-965-6	Sequence 6, Appli
	3	218.6	26.8	846	4	US-09-248-796A-29	Sequence 29, Appl
	4	86.6	10.6	927	4	US-09-248-796A-2490	Sequence 2490, Ap
	5	74.6	9.1	852	4	US-09-734-237B-71	Sequence 71, Appl
	6	61.8	7.6	873	4	US-09-248-796A-2654	Sequence 2654, Ap
	7	55	6.7	336	4	US-09-248-796A-2488	Sequence 2488, Ap
c	8	53	6.5	372	4	US-09-248-796A-6987	Sequence 6987, Ap
	9	53	6.5	774	3	US-09-134-001C-725	Sequence 725, App
c	10	51.8	6.3	825	4	US-09-248-796A-2730	Sequence 2730, Ap
	11	51.2	6.3	852	3	US-09-367-012-2	Sequence 2, Appli
	12	51.2	6.3	852	4	US-09-777-157A-2	Sequence 2, Appli
	13	51.2	6.3	852	4	US-09-734-237B-69	Sequence 69, Appli
	14	51.2	6.3	951	4	US-09-489-0399A-5819	Sequence 5819, Ap
	15	51.2	6.3	1303	3	US-09-367-012-8	Sequence 8, Appli
	16	51.2	6.3	1303	3	US-09-777-157A-8	Sequence 8, Appli
c	17	50.4	6.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	18	49.8	6.1	873	4	US-09-248-796A-2489	Sequence 2489, Ap
	19	48.2	5.9	3993	4	US-09-710-279-3985	Sequence 3985, Ap
c	20	47.4	5.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	21	47	5.8	792	3	US-09-134-001C-1675	Sequence 1675, Ap
	22	45.2	5.5	2774	3	US-09-363-189B-5	Sequence 5, Appli
	23	44.4	5.5	1248	4	US-10-024-806-7	Sequence 7, Appli
	24	43	5.4	879	3	US-09-468-738A-1	Sequence 1, Appli
	25	43	5.3	879	4	US-09-940-019-1	Sequence 1, Appli
	26	43	5.3	879	4	US-09-940-037A-1	Sequence 1, Appli
	27	43	5.3	891	3	US-09-468-738A-22	Sequence 22, Appl
	28	43	5.3	891	3	US-09-940-019-22	Sequence 22, Appl
	29	43	5.3	891	3	US-09-940-037A-22	Sequence 22, Appl
	30	43	5.3	891	3	US-09-468-738A-17	Sequence 17, Appl
	31	43	5.3	1787	3	US-09-940-019-17	Sequence 17, Appl
	32	43	5.3	1787	3	US-09-940-037A-17	Sequence 17, Appl
	33	42.2	5.2	1787	4	US-09-710-279-3157	Sequence 3157, Ap
	34	42.2	5.2	789	4	US-09-710-279-3922	Sequence 3922, Ap
	35	42.2	5.2	3210	4	US-09-710-279-4169	Sequence 4169, Ap
	36	41.4	5.1	3368	4	US-09-710-279-381	Sequence 381, App
	37	41.4	5.1	825	3	US-09-134-001C-1594	Sequence 1594, Ap
	38	41.4	5.1	916	4	US-09-710-279-3426	Sequence 3426, Ap
	39	41.4	5.1	3024	4	US-09-248-796A-148	Sequence 148, App
c	40	41.2	5.0	2472	4	US-09-468-738A-14	Sequence 14, Appl
	41	41.2	5.0	437	3	US-09-940-019-14	Sequence 14, Appl
	42	41	5.0	437	4	US-09-940-037A-14	Sequence 14, Appl
c	43	41	5.0	837	4	US-09-328-352-2671	Sequence 2671, Ap
c	44	40.8	5.0	601	4	US-09-949-016-62984	Sequence 62984, A
c	45	40.8	5.0	50263	4	US-09-949-016-13563	Sequence 13563, A

ALIGNMENTS

RESULT 1
US-08-336-198C-6
; Sequence 6, Application US/08361198C
; Patent No. 5866382
GENERAL INFORMATION:
   APPLICANT: Hallborn, Johan
   APPLICANT: Penttila, Merja
   APPLICANT: Ojamo, Heikki
   APPLICANT: Keranen, Sirkka
   APPLICANT: Hahn-Hagerdal, Barbel
   APPLICANT: Waldfridsson, Mats
   APPLICANT: Airaksinen, Ulla
   TITLE OF INVENTION: XYLOSE UTILIZATION BY RECOMBINANT YEASTS
   NUMBER OF SEQUENCES: 9
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Birch, Stewart, Kolasch & Birch
      STREET: 301 N. Washington St.
      CITY: Falls Church
      STATE: Virginia
      COUNTRY: USA
      ZIP: 22046-3487
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/336,198C
      FILING DATE: 03-NOV-1994
      CLASSIFICATION: 435
   ATTORNEY/AGENT INFORMATION:
      NAME: Murphy Jr., Gerald M.
      REGISTRATION NUMBER: 28,977
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: 703-205-8000
      TELEFAX: 703-205-8050
      TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 696 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
   MOLECULE TYPE: cDNA
   HYPOTHETICAL: NO
   ORIGINAL SOURCE:
      ORGANISM: Pichia stipitis
      STRAIN: CBS-6054
   FEATURE:
      NAME/KEY: CDS

RESULT 2
US-09-184-965-6
Sequence 6, Application US/09184965
Patent No. 6582944
GENERAL INFORMATION:
; APPLICANT: Hallborn, Johan
; APPLICANT: Penttila, Merja
; APPLICANT: Ojamo, Heikki
; APPLICANT: Keranen, Sirkka
; APPLICANT: Hahn-Hagerdal, Barbel
; APPLICANT: Waldridsson, Mats
; APPLICANT: Airaksinen, Ulla
; TITLE OF INVENTION: XYLOSE UTILIZATION BY RECOMBINANT YEASTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-3487
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/184,965
; FILING DATE:
; CLASSIFICATION:
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,198
; FILING DATE: 03-NOV-1994
ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
; LENGTH: 696 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
; ORGANISM: Pichia stipitis
; STRAIN: CBS-6054
FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..693
; OTHER INFORMATION: /standard_name= "xylitol
; OTHER INFORMATION: dehydrogenase"
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: FI 901771
; FILING DATE: 06-APR-1990
US-09-184-965-6

Query Match   29.2%;   Score 238.2; DB 4; Length 696;
Best Local Similarity   64.4%;   Pred. No. 9.4e-65;
Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

```
Qy  227 CCTGTGATATTTCTGATTCTGATACCGTTCACAAGGTGTTTGCTCAAGTTGCTAAGGATT 286
            ||||  ||||  |||||| ||| ||||   ||||||| |||| |||||||||||
Db  107 CCTGCAACATTGGGGATGCTGAGGCAGTAGACCGTTCAGCTCACCATCAACGAACACC 166

Qy  287 TTGGTAAGTTGCCATTGCACTTGGTTAACACAGCTGGTTACTGTGAAAACTTCCCATGTG 346
            |||  ||||| || ||||||| || |||| || ||||| |||||||||||||||
Db  167 ACGGCAAGATCGGTGACTTGTTGATTAACACCCGCTGGATACTGTGAAAAACTTCCCTGCG 226

Qy  347 AAGATTACCCAGCCAAGAACGCTGAGAAGATGGTGAAGGTTAACTTGTTGGGGTTCTTTGT 406
            |||  ||||| || |||| |||||| |||| |||||| |||| || |||||| ||
Db  227 AAACGTACCCGGCTACTAACGCTGAAAAGCATCATGAAGGTGAACGGTTTGGGGTCATTCT 286

Qy  407 ATGTTTCTCAAGCCTTTGCTAAGCCATTGATCAAAGAAGGTATCAAGGGTGCTTCGTGTTG 466
            | ||||| ||| || ||| || |||||| || | ||  |||| ||| || || ||
Db  287 ACGTTTCGCAAATCGTTCGCTAGACAATTGACCAGAA---CAACTTGAGAGAGGCTCTATCA 343

Qy  467 TTTTGATTGGTTCTATGTCTGGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTCTCACA 526
            | ||||| || || | ||||| ||| ||||||||| ||  |||  ||| |||| ||||
Db  344 TCTTGAATTGGCTCAAGTGCTGGAACAATTGTCAACGACCACACCCAAATGTATGTACA 403

Qy  527 ACATGTCCAAGGCTGGTGTTATCCCATTGGCTGAAGACTTTGGCTGTTGTGAAATGGGCTAAGT 586
            ||| |||||||||| || |||| || |||| |  || ||||| || |  |||||||| ||||
Db  404 ACAGTCCAAGGCTGGAGTGATCCACTTGGTCAGATCGTTGGGCTGCGAATGGGCCAAGT 463

Qy  587 ACAACATCAGAGTTAATTCTTTAAACCCAGGTTACATCTACGGGTCCTTTGACCAAGAATG 646
            ||||||||||||| || | |||| |||| || ||||| ||| || ||||| ||||| |
Db  464 ACAACATCAGAGTCAACACCTTATCACCAGGCTATATTTGACTCCTTTAACCAGAAACG 523

Qy  647 TTATCAAATGGTAGCAGAAGAATTGTACACAGATGGATCTCTGGTATCCCACAACAACAAGAA 706
Db  524 TGATTTCTGGCCACACAGAGTGAAAGGAAGCCTGGGAANTCCCATGAAGAGAA 583

Qy  707 TGTCCGGAACCAAAGGAAATACATTGGTGCTGTTCTTTCTGCTTTCTGAAATCTGCTGCTT 766
Db  584 TGGCCGGAACCCAAGGAAATTCGTGGGGGTCCATCTTATATACTTGGCAAGCAACGAGACTGCTTCTT 643

Qy  767 CATACACTACTGGTTGCCAGCTTACTCGTGGTTGATGGTGGTT 805
Db  644 CCTACACTACGGGCCCACAAATTTGGTTGTTGGACGGAGGAT 682
```

; US-08-336-198C-6
; LOCATION: 1..693
; OTHER INFORMATION: /standard_name= "xylitol
; OTHER INFORMATION: dehydrogenase"
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: FI 901771
; FILING DATE: 06-APR-1990

Query Match   29.2%;   Score 238.2; DB 2; Length 696;
Best Local Similarity   64.4%;   Pred. No. 9.4e-65;
Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

```
Qy    647 TTATCAATGGTAACGACGAAATTGTACACAGATGGATCTTCTGTGTATCCCACACAGAAAGAA 706
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    524 TGATTCTGGCCACACAGAGATGAAGGAAGCCTGGGAATCCCATGAAGAGA 583

Qy    707 TGTCCGAACCAAAGGAAATACATTGGTAGTGTGCTTTCTTTCTGAATCTGCTGCTT 766
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    584 TGGCCGAACCCAAAGGAAATTCGTGGGGTCCATCTTATACTTGGCAAAGCGAGACTGCTTCTT 643

Qy    767 CATACAACTACTGGTGTGCCACCTTACTGGTTGATGGTGGTT 805
          ||||||||||||||||||||||||||||||||||||||||||
Db    644 CCTACACTACGGGCCCACCAATTTGGTTGTGGGACGGAGGAT 682
```

```
Query Match           26.8%;  Score 218.6;  DB 4;  Length 846;
Best Local Similarity 62.0%;  Pred. No. 1.6e-58;
Matches  385;  Conservative   0;  Mismatches 224;  Indels  12;  Gaps   2;

Qy     14 TTCCAACTTTTAGATTCGATGGCCACTTAACCATTGTCACAGTGTCGTCTGGTGTGGTTTAG 73
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    218 TCCCAAGCTTCCGTTTGATGGAAAACTAGTCATTAACCGGTGGCTCTGGTGGTTCG 277

Qy     74 CTGAAGCTTTAATCAAAGGGTTTGTTGGGCCTTACGGGTTCTTTCTGGCTGATGATATCG 133
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    278 CTGCGCGTGGTATCAAAGAGCTTTATTAGCCAAAGGTTGGCCAGATTGCTGTGCAGTTCA 337

Qy    134 ACCAAGAAAAGACTGCTGCCAAACAGAGCCGAATTACCACAAATACGCTACTTAGAGAATTCA 193
          ||||                                  ||||||
Db    278 AGGTAAATACGAATCACAATCCGTCAGTGACTGGCTTGGTCATGTAATAATTGGCGATG 457

Qy    194 AGTTGAAGAGAGTTCCAA-----------AGATGGGTTCATATGCCTGTGATAATTTTCTGATT 244
          |||| |||||||||||| 
Db    338 ACTTGGAAAAGAACACAACAAGCTGCTAGAGACCGTCTTACAATGGGGCGGGAGCAAATCA 397

Qy    245 CTGATACCGTTCACAAAGGTGTTTGCCTCAAGGTGTAAGGTTGCCATTGC 304
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    458 CTGAAGCTGTCGACTTGACATTCAAAAGCCATCAACGAACAGGGCAAAATCTCAAGTG 517

Qy    305 ACTTGGTTTAACACAGCTGGTTACTGTGAAGAATTACCCAGCCCAGA 364
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    518 TCTTGGTCAAACTGCCGTTAGCCGTGAAAACTTCCCAGCTGAAGAGTAGTTTCAAGCCATTGA 577

Qy    365 ACGCTGAGAGAGATGGTGAAGGGTTAACTTGTTGGGTTCTTCAAGCCTTTG 424
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    578 ACGCTGAAAACCTTATGAAAGTTAACGGGTTGGTGGTCATTCATTCTCACGTTTCCAAGCTTTTG 637

Qy    425 CTAAGCCATGATTGATCAAAGAGAGGTATCAAAGGGTGCTTCTCGTTGGTTTGATGGTTCTATGT 484
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    638 CTAGACCATTAATTCCAAAA---CAACATGACCGGATCGATCATTTTGATCGGGTCAAATGT 694

Qy    485 CTGTGCCACTTGTCAACCGATCCTCAAAAACCAAGTTGTCTACACACATGTCCAAGGCTGGTG 544
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Query Match           10.6%;  Score 86.6;  DB 4;  Length 927;
Best Local Similarity 52.0%;  Pred. No. 8.1e-17;
Matches  328;  Conservative   0;  Mismatches 279;  Indels  24;  Gaps   5;

Qy    182 CTGAAGAATTGAAGTTGAAAGAGAGAAGTTCCAAAGATGGTTCATATGCCTGTGATATTTCTG 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 CTGAATATTTAACCAAGACTTATGGGGGTTAAATCCAAAGCATACAAAATGTAAGTTACTG 367

Qy    242 ATTCTGATACCGTTCACACAGGTGTTTGCTCAAGTTGCTAAGGATTTTGGTTAAGTTGCCAT 301
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 CTGGGGTCAAACATTTGTCCAAAATCATCACTTGAGGTGTTGAAATGGGGTACCCATTGATA 427

Qy    302 TGCACTTGGTTAACACAGCTGGTTACTG---TGAAAACTTCCATGTGAAGAGATTACCCAG 358
          ||||||||||||||||||||||||||||  
Db    368 ATTTCCCAAGATGTTGAAAAAGTTGTCAAACAAAATGAACTGAATTTCGGTACCATTGATA 427

Qy    359 CCAAGAACGCTGAGAGAGATGGTGAAGGGTTAACTTGTTGGGGTTCTTCTTGTATGTTCCAAG 418
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 TCTTTGTTGCCAATGCTGGTTGTTGCCTGACCAGAGGGCCGGAAATCGATGTGAAATCGAGGGAG 487

Qy    419 CCTTTGCTAAGCCANTTGATCAAAGAGAGGTATCAAAGGGTGCTTCTCGTTGGTTGATGGTT 478
          ||||||||||||||                   
Db    488 TCGACAAAAGGAACAAAAAGTTGTTGAATTAAACAGTGTTTATTAATGTGTTCATG 547

Qy    479 CTATGTCTGGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTCTACACACATGTCCAAGG 538
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 TTGTTGGTCCAATTTTCAGAAAAAAAGGGTA-----AAGGGTCATTCATTTTCACTGCCA 601

Qy    539 CTGGTGTTATCCATTGGGTGTTAAGACTTTGGCTTTGTGAAGTTGACAAAATCAGAG 598
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602 GTAGTCGGCGTTCAAITGTTAAGTTTCCCACAATTTACAAGGCCTGCTGCTAAAG 661

Qy    599 TTAATTCTTTAAAACCCAGGTTACATCTACGGGTCCTTGACCAAGAACTGTTTATCAATGGTA 658
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 CTGGGGGTCAAACATTTGTCCAAAATCATCACCAGTCGTTGAAATGGGGTACCCATTGATA 718

Qy    659 TCAATTCTGTTTCTCCAGTTACATCGCTACTACTCTTAGTGAATTTGCTG------AGAG 718
          ||||||||
Db    719 ACGAAGAATTGTACAACAGATGGATCTCTGGTATCCCACAACAAGGAATGTTCCGAACCAA 769

Qy    719 TCAATTCTGTTTCTCCAGTTACATCGCTACTACTCTTAGTGAATTTGCTG------AGAG 718
          ||||||||
Db    770 ATCCCGGATGTCAAGAGTAAAATGGGTTGCAACTTACACCACTTGGTTAGAGAGAGAGCCAAAACCAA 829
```

```
Qy    719 AGGAATACAATTGGTGCTCTTTTGTACTTGCTTTCTGAATCTGCTTCATCACTACTG 778
          |||||||    |||   ||||||||||||||||||||||||||| ||| |||
Db    830 GAGAGCTTGTTGGGTGCCTACTTGTTATTTGGCTTCCGA---TGCTGCATCTTATACAAACTG 886

Qy    779 GTGCCAGCTTACTTGGTTGATGGTTGGTTTCAC 809
          ||||||||||||||||||||||||||||||||||
Db    887 GTGCTGATCTTGCTGTTGATGGTGGTTACAC 917
```

```
Query Match   9.1%;  Score 74.6;  DB 4;  Length 852;
Best Local Similarity   56.0%;  Pred. No. 4.7e-13;
Matches  209;  Conservative   0;  Mismatches 149;  Indels  15;  Gaps   3;
```

```
Qy    437 TCAAAGAAGGTATCAAGGGTGCTTCTGTTGTTTGATTGGTTCTATGTCTGGTTGCCATTG 496
          ||| || || || |||| ||||| |||  |||||| |  | ||| || || ||||||
Db    485 TCGAAAAGGAAGGTAAGAAAGGCGCCTCTGGTTTCACCGCTTCTTATGTTCTGGTCACATCG 544

Qy    497 TCAACGAATCCTCAAAAACCAAGTTGTCTCAACACATGTCCAAGGCTGGGTTGTTTATCCATTTGG 556
          || |||| ||||| ||||| ||||||||  || |||| || ||||||||| ||||| |||||||
Db    545 TTTAACGTACCGGACCAGTTTCAGGGTTACCTTCCAAGCCTGGGTTGTTCGTCACTTCG 604

Qy    557 CTAAGACTTTGGCTTGTGTGAAATGGGGTAAGTACAACATCAGAGAGTTAATTCTTTAAACCCAG 616
          ||||  ||||||||||||  |||||| |||| || || |||| |||||| || || ||||||||
Db    605 CTAAATCTCTGGCTGTGTAGAAATTCGGCTCCGT---TCGCTCGTGTTAACTTCTGTTTCTCCGG 661

Qy    617 GTTACATCATCTACCGGTCCTTTGACAAAGAAATGTTATCAATGGTAACGAAGAAATTGTACAACA 676
          ||||||||||| || || ||||| |||||| || || ||||||||||||||||| ||||||
Db    662 GCTACAATCAAC------ACCGAAATCTTCGACTTTGTACCGCAGAAAACTCAGAACA 712

Qy    677 GATGGATCTCTGGTATCCCACAACAAAGAATGTCCGACCAAAAGGAAATACAATTGGTGCTG 736
          ||||| ||| || || ||||| ||  |||| ||| ||||| || || ||||||||||
Db    713 AATGGTGGTCTCTGGGCCGTGGCGAAACTGCTGACTGGTTGGTGCTT 772

Qy    737 TTTTGTACTTGCTTTCTGAATCTGCTGTGGTGTGCCAGCTTACTGGTTG 796
          || |||||  ||||| ||  |||||| |||| || |||||| |||||||
Db    773 ACCTGTTTCTGGCTTCTGA---CGCTCGGTTCTTACGGCTACCGGCACTGACTGACTCATCGGTTG 829

Qy    797 ATGGTGGTTTCAC 809
          ||||||||| |||
Db    830 ACGGTGGTTACAC 842
```

```
Query Match   7.6%;  Score 61.8;  DB 4;  Length 873;
Best Local Similarity   51.3%;  Pred. No. 5.3e-09;
Matches  256;  Conservative   0;  Mismatches 222;  Indels  21;  Gaps   4;
```

```
Qy    307 TTGGTTAACAGACTGGTTACTGTGTGAAAAACTTCCCATGTGAAGATTACCCAGCCAAGAAC 366
          |||||||    || ||||| |||||||| |||| ||||  |||||| ||||||||| |||||
Db    364 TTGGTCAATAATGCAAGGAATAACTGAAAAGATAATCCTTTTGAAGATATTCCCACAACAAGATAAG 4423

Qy    367 GCTGAAGAGAGATGGTGAAGGTTAACTTGTTGAGGTATTGTCTTTGTGGGTTTCTTTGTATGTTTTGTCAAGGCCTTTGCT 426
          ||||||| || ||| ||| || ||||| || |||| || || || ||| |||| || |||||||| ||
Db    424 TTTGATGACGTGATAAAAAGTTGATTTGAATATTCCGGTAAAGTTGACTAAAGTATTGGT 483

Qy    427 AAGCCATTGATCAAAGAAGGTATCAAGAGGTGCTTCTGTTGGTTTTGATTGGTTCTATGTCT 486
          ||||||||| || ||| |||| |||| || |||  |||||| |||| |||||| ||||||
Db    484 AGGAAAAATGCTTGAAACAAAATACCAAAGGGTAAGATTGTTTTTTACCGGTTCTTTTATTAT- 542

Qy    487 GGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTCTCAACACATGTCCAAGGCTGTGGTT 546
          ||||| ||||| || ||||| ||||||||| || |||| || ||||||||| ||||||| |||
Db    543 -ATTCCAAGCGCGGGATGGTTGTGCCGACCCC------CTATGCCATCAGTAAAGGGTAAGGTTGCTTTA 594

Qy    547 ATCCATTTGGCTAAGACTTTGGCTGTTGAATGGGGTAAGTACAACATCAGAGTTAATTCT 606
          || ||||| || || ||||| |||||||||| |||||| || |||| |||||| || ||
Db    595 AAACAAATTCACACAAGCAGTATCTTGAATGAATAGTGTCATCAAGAGGGGTATCAGGGTGTCAATTCA 654

Qy    607 TTAAACCCAGGTTACATCTTACGGTCCTTTGACGGTGATCAAACTGTTATCAAATTGTAAGGTAAGGAAA 666
          |||| |||| || ||||| ||  |||||| ||  || || |||  |||| ||||||| ||  ||||
Db    655 ATTGCACCTGGATATATCAAAAACCAATTTGACCCCACACAGGAATGTCCAGACAAGAACAAGAA 714

Qy    667 TTGTACACAGATGGGATCTTCTGGTTATCCCACAACAAACAAAGAGAATGTCCGAACCAAAAGGGAATAC 726
          || |||| || ||| |||| ||| | ||| |||| || || || || ||||| |||| |||||
Db    715 ATTGTTGAT--------TTGAGAATCCCAATGAAAAAGATGGGGTAACCCAAAAGGAACCGGACTTT 765

Qy    727 ATGGTGGTCGTGTTGTTACTTGCTTTCTGAATCTGCTGCTTCATCACTACTGGTTGCCAGC 786
          ||||| ||||||||  ||||| ||||| ||  |||||| || ||| || |||||||||||||
Db    766 ATGGGGCCAATTGTCTACTTTACACATTGA---TGCATCGAAAATATGTTTACTGGTTGACA 822

Qy    787 TTACTGGTTGATGGTGGTT 805
          |||||||||||||||||||
Db    823 TTAATTAGTTGATGGTGGTT 841
```

```
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 2488
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-2488

Query Match             6.7%;   Score 55;  DB 4;  Length 336;
Best Local Similarity  52.9%;   Pred. No. 4.3e-07;
Matches  171; Conservative   0; Mismatches 140; Indels  12; Gaps  2;

Qy    491 CCATTGTCAACGANCCTCAAAAACCAAGTTGTCTTACACANGTCCAAGGCTGGNGTTTATCC 550
          |||||| || | |||||  ||||| |||||||| |  ||||| |   |  || | |||||||
Db     20 CCCGTGGTGAATACCCTCAATTACAAGCTTCCTTATAATGTCTGCTAAAAGCTGCATGGTACTC 79

Qy    551 ATTTGGCTAAGACTTGGCTTGGTGNATGGGGCTAAGTACAACATCAGAGTTAATTCTTTAA 610
          ||  |  || || || |  |  ||| |||| |||| |||| ||| || ||| |||| ||||
Db     80 ATTTAGCCAAAATCATTGAGTGTGGAATGGGCTAGTTTTGGTGTCGTAGAGTNAAATTCAATTT 139

Qy    611 ACCCAGGTTACATCTACGGTCCTTTGACCAAGAATGTTATCAAATGGTAACGAAGAATTGT 670
          |||||||||| || |||  ||||||| |  ||| |  ||| ||   |||  |||||| |||
Db    140 CTCCAGGGTTATATA----------TTGACTGATATTGCTGATTTTGCTGATCCAGAAAATGA 190

Qy    671 ACAAACAGANTGGATCTCTGGTATCCCACAACAAAAAGGAATACATTG 730
          |||||| ||  | || || |||||||  | |  |||||| ||| ||||
Db    191 AAAAGAAANTGGTGGCAATTGACACCTTTGGGAAGAAGGAATTACCACAAAGAATTAGTGG 250

Qy    731 GTGCTGTTTTGTACTTGCTTTTCTGAATCTGCTGCTTCATACACATTCTGGTTGCCAGGTT 790
          ||  || ||| |||  |||  |  |||| |  |||| |  ||| | || || |  || |||
Db    251 GGGCATATTTATACTTGGCCTC---AAATGCATCAACTTATACTACTACGTTCAAATATTG 307

Qy    791 TGdTTGATGGTGGGTTTCACTTCT 813
          ||||| ||||||| |||||||||
Db    308 CTGTTGATGGGGGGTTTATACATGT 330

; SEQ ID NO 2488
Query Match             6.5%;   Score 53;  DB 4;  Length 372;
Best Local Similarity  62.4%;   Pred. No. 2e-06;
Matches  83; Conservative   0; Mismatches  50; Indels  0; Gaps  0;

RESULT 8
US-09-248-796A-6987/c
Sequence 6987, Application US/0924879GA
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6987
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-6987

Qy     14 TTCCAACTTTTAGATTCGANGGCCACTTAACCATTGTCACAGGTGCCCTGGTGGTTTTAG 73
          |||| ||  | |||| ||  ||||||||  | | |||||||| || | |||| |||| |||
Db    205 TCCCAAGCTTCCGTTGGATGGAAAACTAGTCATATTAACCGGTGGCTCTGGGTGGTTTGG 146

Qy     74 CTGAAGCTTTAATCAAGGGTTTGTTGGGCTACGGTTCTGCACATTGCTTTGCGTGGATACG 133
          | |||||| |  |  |||||| |   ||||| ||| | |   |||| ||| || | |||
Db    145 CTGCCGTGTGGTATCAAGAGCTTTATTAGCCAAAGGTGCCGGATGTGTTGCATTAGTCGATATGA 86

Qy    134 ACCAAGAAAAGAC 146
          | |||||| |||
Db     85 ACTTGGAAAGAAC 73
```

```
RESULT 9
US-09-134-001C-725
Sequence 725, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 725
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-725

Query Match             6.5%;   Score 53;  DB 3;  Length 774;
Best Local Similarity  47.8%;   Pred. No. 3e-06;
Matches 223; Conservative   0; Mismatches 235; Indels   9; Gaps  2;

Qy    200 AAGAAGTTCCAAAGANGAGGTTCATAATGCCTGTGATATTTCTGATTCTGATACCGTTCACA 259
          || ||||| |||||| ||| | || |  |||||||| |  |||| |||| |   |||||||
Db    182 AATCAGGTGTTAAAAACACATCATTAAAATTAGATATAATATCATCTAGAATGAAGTAAATC 241

Qy    260 AGGTGTTTGCTTCAAGTTGCTAAAGGATTTTCGTAAGTTGCCATTGCACTTGGTTAACACAG 319
          || |  |  | || ||| |||| |  ||||||| | |||| ||||| ||| || || | ||
Db    242 AAATTGTTGATTAATCTAGTAGGAATATGGAAAAAATAGATATTC---TTGTCAATAACG 298

Qy    320 CTGGTTACTGTGAAAACTTCCCATGTGAAGATTACCACGCCAGAAACGCTGAGAAGGATGG 379
          || ||||| |    ||| |||| | ||  |   | ||  ||| |||||  | || |||| |
Db    299 CTAGTTATCTCAATTCAAGATGATACTGAAAATATTTCGTATGAAGAATGGTTAAAAGAAA 358

Qy    380 TGAAGGTTAACTTGTTGGGGTTCTTTTGTATGGTTTCTCAAGCCTTTGCCTAAGCCATTGATCA 439
          ||| || || |||||| | |||| |||| |  ||||| ||| |||||| ||| |||| |||
Db    359 TCAATTTAAGCTTAAAATGGAGCAATTCTCGTTGGGCAAACAGTTAGGACGTCAAAATGATTG 418

Qy    440 AAGAAGGTATCAAGGGGTGCTTCTGTTGTTGTTTTGATTGGTTCTATGTCTGGTCCCATGTCA 499
          || ||| | ||| ||| |||  ||| || | |||  | |  |||  ||||  |||  |||||
Db    419 AGAACGGCA------GTGGTTCGATGATTAATGTCTTCTGAGTATTAGGAATTGCTA 472

Qy    500 ACGATCCTCAAAACCAAGTTGTCTACAACATGTCCAAGGCTGGGTGTTATCCATTTGGCTA 559
          || || | ||||  ||||| | || |  ||||| | |||  ||| |   || |||| |
Db    473 ATAAAACGACAAGACCAGAGCTCTTATGAAACGTCTAAAGCCGGTGTGACAGATGCTGACTA 532

Qy    560 AGACTTTGGCTTGTGAATGGGCTAAGTACAACATCAGAGTTAATTCTTTAAACCCAGGTT 619
          || |||| || | || ||||| ||| |||| ||| ||||| ||||| ||||| ||||||
Db    533 AAAGCTTAGCTAGAGAATGGTCTAGATATGGGAATTAAAGTAAATGCAATTGCCCCTGGTT 592

Qy    620 ACATCTACGGTCCTTTGACCAAGAATGTTATCAAATGGTAACGAAGAA 666
          |||||  ||| |||| ||||| |  |  | |||||||| | |||||||
Db    593 ACATGAGAACAAATCGAAAACGGAAAAAGATACTAAAATGATAATACTGAA 639

RESULT 10
US-09-248-796A-2730
Sequence 2730, Application US/0924879GA
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
```

```
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 2730
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-2730

Query Match      6.3%;  Score 51.8;  DB 4;  Length 825;
Best Local Similarity   45.9%;  Pred. No. 7.3e-06;
Matches 299;  Conservative  0;  Mismatches 337;  Indels 15;  Gaps  3;

Qy     22 TTTAGATTCGATGGCCCACTTAACCATTGTCACAGGTGCCGTGGTGGTTTAGCTGAAGCT 81
             |    | |  | ||  | |      |    |   |    |    | |   || |  |||
Db     25 TTCAATGTTAATGGGAAAATTGCTGTGTTGTTAACTGGTGGTACTCGGTGTTGGGTTTATAC 84

Qy     82 TTAATCAAGGGGTTTGTTGGGCCTACGGGTTCTGACATTGCTTTGTTGATATCGACCAAGAA 141
             ||| || | || || |    | |  | |  | |   || ||   || | |  | || |
Db     85 TGTGCTGAGGCCTTTGTTTATTGAACGGGTCATTACATCAGAAGAAAGCA 144

Qy    142 AAGACTGCTGCCAAACAAGCCGAATACCACAAATACGGCTACTGAAGAATTGAAGTTGAAA 201
             | ||| |  |  || |  | | ||  | |  || |  |    | | ||  || ||  |
Db    145 AAAGCTTGTGAACAGCTCAAAAAGTATTTGGAAAAACTTGCCAAAGACAACACAAGA- 203

Qy    202 GAAGTTCCAAAGATGGGTTCATATGCCTGTGATATTTTCTGATTCTGATACCGTTCACAAG 261
             | | |  |||  |  ||  | |  |  | | || |  |   | | ||  |  | | ||
Db    204 -----TTGTAAAATATTTCATATCCTCCAGATATTGCTGTGTTGAAGAAGAAGTGAAAAA 258

Qy    262 GTGTTTGCTTCAAGTTGCTTAAGGATTTGGTAAGTTGCCATTGCCACTTGGTTAACACAGCT 321
             | | |  |  |||  | ||    || |  |  | | | | |  | | |  |||  |  |
Db    259 TTCTACGCCGAAATAGCCAAACAAGTCGACAAAGTTGACAT---TTTGGTCGCCACGCT 315

Qy    322 GGTTACTGTGAAAACTTCCATGTGACGATTACCCAGCCAAGAACGCTGAGAGAGATGGTG 381
             ||| || |    | |  |  |  |    |||  || | | ||  | |||| | ||  |
Db    316 GGTGCTAGTTGGGGTCGCTCCGTTAGAGACACCCAGTTTCAGCAGTCAAAAAAGTCCTC 375

Qy    382 AAGGTTAACTTGGTGGGGTTCTTTGTATGTTTCTCAAGCCTTTGCTAAGCCATTGATCAAA 441
             || ||| | ||| | | | ||  |  || | | || ||  || ||   | || ||  |
Db    376 AACTTGAATGTTGTTGCGTCTACCACACATCAAGTTAATTTTACTCCTTTATTGGAAAA 435

Qy    442 GAAGGTATCAAGGGTGCTTCGTGTTGTTTTGATTGGT------TCTATGTCTCGGTGCCATT 495
             | || ||||  | | | || ||| ||  | || ||       | | ||||| |  | ||
Db    436 GCTCGGTACCAAAGAAGATCCTTCTAGATTGATCCTTAAGTCCTCTCGTCCATTAGTTATTCG 495

Qy    496 GTCAACGATCCTCAAAAACCAAGTTGTCTACAACATGTCCAAGGCTGGTTGTTATCCATTTG 555
             || || || || ||   ||| |  | || | | ||  |  ||||| || | | | || ||
Db    496 ACTAACGGATATGGGTCGGTGTGTATACGGTTATTTGTTCCAAGGCTGCAGTATCTCATTTG 555

Qy    556 GCTAAGACTTTGGCTGTGAATGGGCGTAAGTACAACATCACAGAGTTAATTCTTTAAACCCA 615
             || || ||||| |  | || | | | ||| | ||  | ||    | ||   ||  |  |
Db    556 GGAAGAAACTTGTCTGTTCAATTTGCCCAAGACATAATCAAGTTAATTCAAATTGCCCA 615

Qy    616 GGTTACATCTACGGTCCTTTGACCAAGAATGTTATCAATGGTAACGAAGAA 666
             || ||| |  | | | || |  | ||||  |  | | | | ||| |||
Db    616 GGTTTTTTTCCCATCGAAAATGCCCAATGGTTTGATCGAAGCTGCTGGTGAA 666
```

RESULT 11

```
US-09-367-012-2
Sequence 2, Application US/0967012
Patent No. 6218156
GENERAL INFORMATION:
; APPLICANT: Yasohara, Yoshihiko
; APPLICANT: Kizaki, No. 6218156iyuki
; APPLICANT: Hasegawa, Junzo
; APPLICANT: Wada, Masaru
; APPLICANT: Shimizu, Sakayu
; APPLICANT: Kataoka, Michihiko
; APPLICANT: Yamamoto, Kazuhiko
; APPLICANT: Kawabata, Hiroshi
; APPLICANT: Kita, Keiko
US-09-367-012-2

Query Match      6.3%;  Score 51.2;  DB 3;  Length 852;
Best Local Similarity   58.6%;  Pred. No. 1.2e-05;
Matches  89;  Conservative  0;  Mismatches  63;  Indels  0;  Gaps  0;

Qy    437 TCAAAGAAGGTATCAAGGGTGCTTCGTGTTGTTTTGATTGGTTCTATGTCTCGGTGCCATTGG 496
            | | | ||| |||  | | | |||  | | | |  | | | || | ||| | |  ||  ||
Db    485 TCGAGAAGGAGGGGCAAGAAGGGCGCCCCTTGTTGTTCACGGCGCCTCCATAGTCTGGCCACATTG 544

Qy    497 TCAACGATCCTCAAAAACCAAGTTGTCTACAACATGTCCAAGGCTGGTTGTTATCCATTTGG 556
            || || |  ||   |   | ||  | || | | | | | |||| | ||  |  |  ||  |
Db    545 TGAACGTGCCCCAGTTCCAGGCCACAGTACAACGCGGCCAAGGCTGGCGTGCGCCCACTTCG 604
```

RESULT 12

```
US-09-777-157A-2
Sequence 2, Application US/09777157A
Patent No. 6448052
GENERAL INFORMATION:
; APPLICANT: Yasohara, Yoshihiko
; APPLICANT: Kizaki, No. 6448052iyuki
; APPLICANT: Hasegawa, Junzo
; APPLICANT: Wada, Masaru
; APPLICANT: Shimizu, Sakayu
; APPLICANT: Kataoka, Michihiko
; APPLICANT: Yamamoto, Kazuhiko
; APPLICANT: Kawabata, Hiroshi
; APPLICANT: Kita, Keiko
; TITLE OF INVENTION: Carbonyl Reductase Enzyme and Methods for its Use
; FILE REFERENCE: 068383.O110
; CURRENT APPLICATION NUMBER: US/09/777,157A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/367,012
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Candida magnoliae
US-09-777-157A-2

Query Match      6.3%;  Score 51.2;  DB 3;  Length 852;
Best Local Similarity   58.6%;  Pred. No. 1.2e-05;
Matches  89;  Conservative  0;  Mismatches  63;  Indels  0;  Gaps  0;

Qy    437 TCAAAGAAGGTATCAAGGGTGCTTCGTGTTGTTTTGATTGGTTCTATGTCTCGGTGCCATTGG 496
            | | | ||| |||  | | | |||  | | | |  | | | || | ||| | |  ||  ||
Db    485 TCGAGAAGGAGGGGCAAGAAGGGCGCCCCTTGTTGTTCACGGCGCCTCCATAGTCTGGCCACATTG 544

Qy    497 TCAACGATCCTCAAAAACCAAGTTGTCTACAACATGTCCAAGGCTGGTTGTTATCCATTTGG 556
            || || |  ||   |   | ||  | || | | | | | |||| | ||  |  |  ||  |
Db    545 TGAACGTGCCCCAGTTCCAGGCCACAGTACAACGCGGCCAAGGCTGGCGTGCGCCCACTTCG 604

Qy    557 CTAAGACTTTGGCTGTTGAATGGGCTAAGTAC 588
            || || |||||  | | |||||| | ||||||
Db    557 CTAAGACTTTGGCTGTTGAATGGGCTAAGTAC 588
```

; TITLE OF INVENTION: No. 6218156e1 carbonyl reductase, gene coding same, and method
; TITLE OF INVENTION: for using such reductase and gene
; FILE REFERENCE: S7203OPCUS
; CURRENT APPLICATION NUMBER: US/09/367,012
; CURRENT FILING DATE: 1999-11-24
; EARLIER APPLICATION NUMBER: PCT/JP97/03051
; EARLIER FILING DATE: 1997-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Candida magnoliae
US-09-367-012-2

```
Db   605  CGAAGTCGCTGGGCCGTCGAGTTCGCGCGCGTTC  636
          ||||   |||||  |||| | || ||| || |||
```

RESULT 13
US-09-734-237B-69
; Sequence 69, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583.40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 69
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Candida magnoliae
US-09-734-237B-69

Query Match   6.3%; Score 51.2; DB 4; Length 852;
Best Local Similarity   58.6%;
Matches   89; Conservative   0; Mismatches 63; Indels   0; Gaps   0;

```
Qy   437  TCAAAGAAGGTATCAAAGGGTGCTTCTGTTGTTTTGATTGGTTCTATGTCTGGTGCCATTG  496
          ||| ||||||| ||| |||| |||||| | || ||| |||| || |||| ||| |||||||
Db   485  TCGAGAAGAGGGCAAGAAGGGCCGCCCTTGTGTTCACGGCGTCATGTCTGGTGCCACATTG  544

Qy   497  TCAACGATCCTCAAAACCAAGTTGTCTACAACACATGTCCAAGGCTGGGTGTTATCCATTTGG  556
          ||  || | ||| | ||||| ||  |||||| |||| ||| || |||||| || |||| |||
Db   545  TGAACGTGCCCCAGTTCCAGGCCACGTACAACGCGCCAAGGCTGGGCGTCGCGCCACTTCG   604

Qy   557  CTAAGACTTTGGCTTGTGTGAATGGGCTAAGTAC  588
          |||| ||||||| ||| | ||| ||||  |||||
Db   605  CGAAGTCGCTGGGCCGTCGAGTTCGCGCGCGTTC  636
```

Db   655 CAGGCGCCATTACAACTGCTCTCCAAAGCCGCGTGGTGATCCATCTGTGTCGAAAAGCCTTGTGCGATG 714
          || | ||||| || || | || ||| | || | | || | || | || ||| || |

```
Qy   574  GAATGGGGCTAAGTACACAAACATCAGAGTTAATTCTTTAAACCCAGGTTACATCTACGGTCCT  633
Db   715  GAGTGGGGTGGGGCAAAGGAGTCCGCCGTCAATTCCATCAGTGCCGGGATAATACCGCCACGGCG  774
Qy   634  TTGA 637
          ||||
Db   775  ATGA 778
```

RESULT 14
US-09-489-039A-5819
; Sequence 5819, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5819
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5819

Query Match   6.3%; Score 51.2; DB 4; Length 951;
Best Local Similarity   54.9%; Pred. No. 1.2e-05;
Matches 101; Conservative   0; Mismatches 83; Indels   0; Gaps   0;

```
Qy   454  GGTGCTTCTGTTGTTTTGATTGGTTCTATGTCTGGTGCCATTGTCAACGATCCTCAAAAC  513
          |  |  | ||| ||| || || | ||| | |||   |||| |||||  |  |  ||| ||
Db   595  GGCGGGTCGGATTATCAATATCGCCTCGATGTCGGGGGATTATTGTTAACCGGCGCCTCGAC  654
Qy   514  CAAGTTGTCTACAACATGTCCAAGGCTGGGTGTTATCCATTTGGCTAAGACTTTGGCTTGT  573
```

RESULT 15
US-09-367-012-8
; Sequence 8, Application US/09367012
; Patent No. 6218156
; GENERAL INFORMATION:
; APPLICANT: Yasohara, Yoshihiko
; APPLICANT: Kizaki, No. 6218156iyuki
; APPLICANT: Hasegawa, Junzo
; APPLICANT: Wada, Masaru
; APPLICANT: Shimizu, Sakayu
; APPLICANT: Kataoka, Michihiko
; APPLICANT: Yamamoto, Kazuhiko
; APPLICANT: Kawabata, Hiroshi
; APPLICANT: Kita, Keiko
; TITLE OF INVENTION: No. 6218156e1 carbonyl reductase, gene coding same, and method
; TITLE OF INVENTION: for using such reductase and gene
; FILE REFERENCE: S72030PCUS
; CURRENT APPLICATION NUMBER: US/09/367,012
; CURRENT FILING DATE: 1999-11-24
; EARLIER APPLICATION NUMBER: PCT/JP97/03051
; EARLIER FILING DATE: 1997-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1303
; TYPE: DNA
; ORGANISM: Candida magnoliae
US-09-367-012-8

Query Match   6.3%; Score 51.2; DB 3; Length 1303;
Best Local Similarity   58.6%; Pred. No. 1.5e-05;
Matches   89; Conservative   0; Mismatches 63; Indels   0; Gaps   0;

```
Qy   437  TCAAAGAAGGTATCAAAGGGTGCTTCTGTTGTTTTGATTGGTTCTATGTCTGGTGCCATTG  496
          ||| ||||||| ||| |||| |||||| | || ||| |||| || |||| ||| |||||||
Db   643  TCGAGAAGAGGGCAAGAAGGGCCGCCCTTGTGTTCACGGCGTCATGTCTGGTGCCACATTG  702
Qy   497  TCAACGATCCTCAAAACCAAGTTGTCTACAACACATGTCCAAGGCTGGGTGTTATCCATTTGG  556
          ||  || | ||| | ||||| ||  |||||| |||| ||| || |||||| || |||| |||
Db   703  TGAACGTGCCCCAGTTCCAGGCCACGTACAACGCGCCAAGGCTGGGCGTCGCGCCACTTCG   762
Qy   557  CTAAGACTTTGGCTTGTGTGAATGGGCTAAGTAC  588
          |||| ||||||| ||| | ||| ||||  |||||
Db   763  CGAAGTCGCTGGGCCGTCGAGTTCGCGCGCGTTC  794
```

Search completed: March 2, 2005, 16:57:18
Job time : 195 secs

This Page Blank (Jspio)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: March 2, 2005, 04:46:41 ; Search time 560 Seconds
     (without alignments)
     8625.909 Million cell updates/sec

Title: US-10-720-018-1
Perfect score: 816
Sequence: 1 atgactgactacattccaac.........atggtggttcaacttcttgg 816
Scoring table: IDENTITY_NUC
     Gapop 10.0 , Gapext 1.0
Searched: 4390206 seqs, 2559870667 residues
Total number of hits satisfying chosen parameters: 8780412
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
     Maximum Match 100%
     listing first 45 summaries

Database :   N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	816	100.0	816	12	ADQ94601	Adq94601 Ambrosioz
2	99.4	12.2	3203	3	AAF07518	Aaf07518 Fusarium
3	88.6	10.9	687	3	AAF12409	Aaf12409 Aspergill
4	81.8	10.0	637	13	ADS61049	Ads61049 Bacterial
5	80.6	9.9	735	13	ADS48561	Ads48561 Bacterial
6	74.6	9.1	852	5	AAH74598	Aah74598 Synthetic
7	74.2	9.1	1146	3	AAF12302	Aaf12302 Aspergill
8	67.6	8.3	783	13	ADS47380	Ads47380 Bacterial
9	66.4	8.1	789	12	ADN62512	Adn62512 A. thalia
10	63.6	7.8	774	13	ADS48180	Ads48180 Bacterial
11	60.6	7.4	636	3	AAF11343	Aaf11343 Aspergill
12	59.8	7.3	822	6	AAL57425	Aal57425 Datura st
13	59	7.2	774	3	AAZ46762	Aaz46762 Bacillus
14	59	7.2	2255	3	AAZ46763	Aaz46763 Bacillus
15	56.4	6.9	798	3	AAC42189	Aac42189 Arabidops
16	56.4	6.9	825	6	AAL57426	Aal57426 Hyoscyamu
17	56.2	6.9	789	3	AAC43167	Aac43167 Arabidops
18	56.2	6.9	789	3	ABZ13143	Abz13143 Arabidops
19	56.2	6.9	789	12	ADN62514	Adn62514 A. thalia
20	55.2	6.8	732	13	ADS63286	Ads63286 Bacterial
21	55.2	6.8	732	13	ADS63285	Ads63285 Bacterial
22	55.2	6.8	943	12	ADN62526	Adn62526 A. thalia
23	55	6.7	949	6	ABN98571	Abn98571 Arabidops
24	54.4	6.7	783	3	AAC43166	Aac43166 Arabidops
25	54.4	6.7	783	6	ABZ13220	Abz13220 Arabidops
26	53.8	6.6	768	13	ADS48201	Ads48201 Bacterial
27	53.8	6.6	1003	12	ADN62520	Adn62520 A. thalia
28	53.2	6.5	852	12	ADJ98111	Adj98111 Ketoreduc
29	53.2	6.5	852	12	ADJ98097	Adj98097 Ketoreduc
30	53.2	6.5	855	12	ADJ98103	Adj98103 Ketoreduc
31	53.2	6.5	855	12	ADJ98105	Adj98105 Ketoreduc
32	53	6.5	774	6	ABN91262	Abn91262 Staphyloc
33	53	6.5	774	13	ADS01477	Ads01477 Staphyloc
34	52.8	6.5	819	12	ADJ12063	Adj12063 Maize cDN
35	52.8	6.5	1366	12	ADN62538	Adn62538 A. thalia
36	52.4	6.4	656	6	ABQ65920	Abq65920 Arabidops
37	52.4	6.4	789	12	ADN62500	Adn62500 A. thalia
38	52.2	6.4	852	12	ADJ98109	Adj98109 Ketoreduc
39	51.6	6.3	852	12	ADJ98113	Adj98113 Ketoreduc
40	51.6	6.3	852	12	ADJ98115	Adj98115 Ketoreduc
41	51.6	6.3	855	12	ADJ98101	Adj98101 Ketoreduc
42	51.4	6.3	849	12	ADZ32459	Adz32459 Candida a
43	51.2	6.3	711	11	ADJ12134	Adj12134 Maize cDN
44	51.2	6.3	810	6	ABZ13885	Abz13885 Arabidops
45	51.2	6.3	810	8	ABZ42061	Abz42061 Arabidops

ALIGNMENTS

RESULT 1
ADQ94601

ID   ADQ94601 standard; cDNA; 816 BP.
XX
AC   ADQ94601;
XX
DT   23-SEP-2004 (first entry)
XX
DE   Ambrosiozyma monospora NADH dependent L-xylulose reductase cDNA.
XX
KW   NADH dependent L-xylulose reductase; fermentation; biomass; ALX1; gene;
KW   ss.
XX
OS   Ambrosiozyma monospora.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..816
FT   /*tag=   a
FT   /product= "Ambrosiozyma monospora NADH dependent L-
FT   xylulose reductase"
FT   /partial
FT   /note= "No stop codon"
FT   /EC_number= "EC 1.1.1.10"
XX
PN   US2004132074-A1.
XX
PD   08-JUL-2004.
XX
PF   24-NOV-2003; 2003US-00720018.
XX
PR   16-FEB-2001; 2001FI-00000308.
PR   15-FEB-2002; 2002WO-FI000125.
PR   10-MAR-2003; 2003US-00257821.
PR   12-SEP-2003; 2003FI-00001307.
XX
PA   (VALW) VALTION TEKNILLINEN TUTKIMUSKESKUS.
XX
PI   Verho R, Richard P, Penttila M;
XX
DR   WPI; 2004-517001/49.
DR   P-PSDB; ADQ94602.
XX
PT   New DNA molecule encoding NADH dependent L-xylulose reductase, useful in

```
PT   producing fermentation products or in conversion of cheap biomass to
PT   useful products.
XX
PS   Claim 6; SEQ ID NO 1; 15pp; English.
XX
CC   The present invention relates to a NADH dependent L-xylulose reductase
CC   enzyme protein and its encoding polynucleotide. The invention is useful
CC   in producing fermentation products or in conversion of cheap biomass to
CC   useful products. The present sequence is an Ambrosiozyma monospora NRRL Y
CC   -1484 NADH dependent L-xylulose reductase cDNA, ALX1 (EC 1.1.1.10).
XX
SQ   Sequence 816 BP; 217 A; 159 C; 182 G; 258 T; 0 U; 0 Other;

Query Match             100.0%;  Score 816;  DB 12;  Length 816;
Best Local Similarity   100.0%;  Pred. No. 2.7e-227;
Matches  816;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACTGACTACATTCCAACTTTTAGATTCGATGGCCACTTAACCATTGTCACAGGTGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACTGACTACATTCCAACTTTTAGATTCGATGGCCACTTAACCATTGTCACAGGTGCC 60

Qy         61 TGTGGTGGTTTAGCTGAAGCTTTAATCAAGGGGTTTGTTGGCCTACGGTTCTGACATTGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTGGTGGTTTAGCTGAAGCTTTAATCAAGGGGTTTGTTGGCCTACGGTTCTGACATTGCT 120

Qy        121 TTGCTTGATATCGACCAAGAAAAAGAACTGCTGCCAAAACACCACAAAATACCCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTGCTTGATATCGACCAAGAAAAAGAACTGCTGCCAAAACACCACAAAATACCCT 180

Qy        181 ACTGAAGAAATTGAAGTTGAAAGAAGTTCCAAAGATGGGTTCATATGCCTGTGATATTTCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTGAAGAAATTGAAGTTGAAAGAAGTTCCAAAGATGGGTTCATATGCCTGTGATATTTCT 240

Qy        241 GATTCTGATACCGTTCACAAGGTGTTGCTCAAGGTTGCTAAGGATTTTGGTAAGTTGCCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATTCTGATACCGTTCACAAGGTGTTGCTCAAGGTTGCTAAGGATTTTGGTAAGTTGCCA 300

Qy        301 TTGCACTTGGTTAACACAGCCTGGTTACTGTGAAAACTTCCCATGTGAAGATTACCCAGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTGCACTTGGTTAACACAGCCTGGTTACTGTGAAAACTTCCCATGTGAAGATTACCCAGCC 360

Qy        361 AAGAACCGGTGAGAAGATGGTGAAGGTTAACTTGTTGGGGTTCTTTGTATGNTTTCTCAAGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGAACCGGTGAGAAGATGGTGAAGGTTAACTTGTTGGGGTTCTTTGTATGNTTTCTCAAGCC 420

Qy        421 TTTTGCTAAGCCATTGATCAAAAGAAGGTATCAAAGGGTGCTTCTCGTTGTTGTTTGATTGGTTCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTTTGCTAAGCCATTGATCAAAAGAAGGTATCAAAGGGTGCTTCTCGTTGTTGTTTGATTGGTTCT 480

Qy        481 ATGTCTCGGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTTGTCTACACAGTCCAAGGCT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGTCTCGGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTTGTCTACACAGTCCAAGGCT 540

Qy        541 GGTGTTATCCATTGGCTAAGACTTTGGCTAAGACTGTGTTGAAGTACAACATCAGAGTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGTGTTATCCATTGGCTAAGACTTTGGCTAAGACTGTGTTGAAGTACAACATCAGAGTT 600

Qy        601 AATTCTTTAAACCCAGGTTACATCTACGGTGTCCTTTGACCAGNAATGTTATCAATGGGTAAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AATTCTTTAAACCCAGGTTACATCTACGGTGTCCTTTGACCAGNAATGTTATCAATGGGTAAC 660

Qy        661 GAAGAAATTGTACAAACAGATGGATCTCTGGTATCCCACAACAAAGAANTCCGAACCAAAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGAAATTGTACAAACAGATGGATCTCTGGTATCCCACAACAAAGAANTCCGAACCAAAG 720

Qy        721 GAATACAATTGGGTGGCTGTTTGTACTTGCTTCTGAATCGTGCTAATACTACACTACTGGT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAATACAATTGGGTGGCTGTTTGTACTTGCTTCTGAATCGTGCTAATACTACACTACTGGT 780

Qy        781 GCCAGCTTACTGGTTGATGGTGGTTTCACTTCTTGG 816
              ||||||||||||||||||||||||||||||||||||
Db        781 GCCAGCTTACTGGTTGATGGTGGTTTCACTTCTTGG 816
```

RESULT 2
AAF07518
ID   AAF07518 standard; cDNA; 3203 BP.
XX
AC   AAF07518;
XX
DT   13-MAR-2001 (first entry)
XX
DE   Fusarium venenatum EST SEQ ID NO:41.
XX
KW   Multiple gene expression; filamentous fungal cell; EST;
KW   expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW   Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW   culture condition; environmental stress; spore morphogenesis;
KW   metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS   Fusarium venenatum.
XX
PN   WO200056762-A2.
XX
PD   28-SEP-2000.
XX
PF   22-MAR-2000; 2000WO-US007781.
XX
PR   22-MAR-1999;   99US-00273623.
XX
PA   (NOVO) NOVO NORDISK BIOTECH INC.
PA   (NOVO) NOVO NORDISK AS.
XX
PI   Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
DR   WPI; 2000-594572/56.
XX
PS   Claim 86; Page 383-384; 3161pp; English.
XX
CC   The present invention describes a method for monitoring differential
CC   expression of genes in a first filamentous fungal (FF) cell relative to
CC   expression of the same genes in one or more second filamentous fungal
CC   cells. The method uses fluorescence-labeled nucleic acids isolated from
CC   the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC   are used in the methods for monitoring differential expression of genes
CC   in a first filamentous fungal (FF) cell relative to expression of the
CC   same genes in one or more second filamentous fungal cells. Monitoring the
CC   global expression of genes from FF cells allows the production potential
CC   of the microorganisms to be improved. New genes may be discovered
CC   possible functions of unknown open reading frames can be identified and
CC   of genes can be used to study how FF cells adapt to changes in culture
CC   conditions, environmental stress, spore morphogenesis, recombination,
CC   metabolic or catabolic pathway engineering. Using ESTs provides several
CC   advantages over genomic or random cDNA clones including elimination of
CC   redundancy as one spot on an array equals one gene or open reading frame,
CC   and organisation of the microarrays based on function of the gene
CC   products to facilitate analysis of the results. AAF07478 to AAF11247
CC   represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents
CC   ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from
CC   Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from
CC   Trichoderma reesei, which are all specifically claimed in the present
CC   invention
XX
SQ   Sequence 3203 BP; 770 A; 861 C; 815 T; 0 U; 7 Other;

Query Match   12.2%; Score 99.4; DB 3; Length 3203;
Best Local Similarity   52.7%; Pred. No. 4.5e-18;
Matches 265; Conservative   0; Mismatches 231; Indels   7; Gaps   2;

```
Qy        308 TGGTTAACACAGCTGGTTACTGTGAAAACTTCCCATGTGAAGATTACCCAGCCAAGAACG 367
              |||  |||||| ||  |  ||| |  |  | || || |  |    | || ||||||||| |
```

```
Db  2614 TAGTCACGTTCGGCTTGGCTTCACCGAGAGAATTTCGAGGCAATCAACTATCCCATTCGATCGTA 2673
Qy   368 CTGAGAAGATGGTGAAGGTTAACTTGTTGTTGGGTTCTTTGTATGTTTCTCAAGGCCTTTGCTA 427
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2674 TGCCGCAAGTTGTGTGGGGGTGTTAAATGTTGACGGTACTTACCTTCTTTGCAGTTGCAGTTGCCA 2733
Qy   428 AGCCATTGAATCAAAGAAGGTATCAAGGGTGCTTCGTGTTGTTTGATTGGTTCTATGTCTCG 487
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2734 A---ACATCTCATCGAACCGCCAGGTACCTGTAGTATCGTGGTTATTGAAGCANTGTCTCG 2790
Qy   488 GTGGCCATTGTCAACGATCCTCAAAAACCAAGTTGTCTACAACATGTCCAAGGCTGGGTGTTA 547
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2791 GTGCTATTGTCAATGTTCCACAGCGCCACATATAAACGCGGCCAAGGCAGCTGTTC 2850
Qy   548 TCCATTTGGCTAAGACTTTGGCTTGTGTGAAGGGGCTAAGTACAACATCAGAGTTATTCTT 607
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2851 GTCACCTGGCTGCTTCCCTCGCAGTAGAGTGGGGTCACCGTGGAATCCGAGTCAACTGTA 2910
Qy   608 TAAACCCAGGTTACATCTACGGTGTCCTTTGACCAAGAATGTTATCAATGTGAAGAGAAT 667
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2911 TCTCTCCTGGCTACATGTTGACTGCTCTAACGCACAGAAGAGTTCTGAAGATAACCCAGATA 2970
Qy   668 TGTACAAACAGATGGAATCTCTGGTATCCACACAACAAAGAATGTCCGAACCAAAGGAATACA 727
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2971 TCGAGGAGGACCTGGACACATCTCTTATTCCTCAGGGACTACCTCAAGATCTGA 3030
Qy   728 TTGCGTGCTGTTTTGTACTTGCTTTCTGAATCTGCTGCTCATACACTACTTGGTCGTGCCAGCT 787
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3031 TGGGGTCCTGT----GACCTTTTTGTTGCAGATGGGCTTCTTCTTACATGGGCGGCAGATC 3086
Qy   788 TACTTGGTTGGATGGTTGATGGTGGTTTCACT 810
         ||||||||||||||||||||||||||||||||
Db  3087 TTAGATAGATGGAGGAATAYACT 3109
```

```
RESULT 3

AAP12409

ID   AAP12409 standard; cDNA; 687 BP.
XX
AC   AAP12409;
XX
DT   13-MAR-2001 (first entry)
XX
DE   Aspergillus oryzae EST SEQ ID NO:4932.
XX
KW   Multiple gene expression; filamentous fungal cell; EST;
KW   expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW   Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW   culture condition; environmental stress; spore morphogenesis;
KW   metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS   Aspergillus oryzae.
XX
PN   WO200056762-A2.
XX
PD   28-SEP-2000.
XX
PF   22-MAR-2000; 2000WO-US007781.
XX
PR   22-MAR-1999; 99US-00273623.
XX
PA   (NOVO ) NOVO NORDISK BIOTECH INC.
PA   (NOVO ) NOVO NORDISK AS.
XX
PI   Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX
DR   WPI; 2000-594572/56.
XX
PT   Monitoring differential expression of genes in filamentous fungal cells
PT   uses fluorescence-labeled nucleic acids isolated from the cells and a
PT   substrate of expressed sequence tags.
XX
PS   Claim 88; Page 2078; 3161pp; English.
XX
CC   The present invention describes a method for monitoring differential
CC   expression of genes in a first filamentous fungal (FF) cell relative to
CC   expression of the same genes in one or more second filamentous fungal
CC   cells. The method uses fluorescence-labelled nucleic acids isolated from
CC   the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC   are used in the methods for monitoring differential expression of genes
CC   in a first filamentous fungal (FF) cell relative to expression of the
CC   same genes in one or more second filamentous fungal cells. Monitoring the
CC   global expression of genes from FF cells allows the production potential
CC   of the microorganisms to be improved. New genes may be discovered and
CC   possible functions of unknown open reading frames can be identified and
CC   gene copy number variation and stability can be monitored. The expression
CC   of genes can be used to study how FF cells adapt to changes in culture
CC   conditions, environmental stress, spore morphogenesis, recombination,
CC   metabolic or catabolic pathway engineering. Using ESTs provides several
CC   advantages over genomic or random cDNA clones including elimination of
CC   redundancy as one spot on an array equals one gene or open reading frame,
CC   and organisation of the microarrays based on function of the gene
CC   products to facilitate analysis of the results. AAP07478 to AAP11247
CC   represents ESTs from Fusarium venenatum; AAP11248 to AAP11853 represents
CC   ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from
CC   Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from
CC   Trichoderma reesei, which are all specifically claimed in the present
CC   invention
XX
SQ   Sequence 687 BP; 152 A; 210 C; 175 G; 147 T; 0 U; 3 Other;
XX
```

```
Query Match             10.9%;  Score 88.6;  DB 3;  Length 687;
Best Local Similarity   49.8%;  Pred. No. 3.3e-15;
Matches  279;  Conservative   0;  Mismatches  275;  Indels    6;  Gaps    2;
```

```
Qy   196 TTGAAGAAGTTCCAAAAGATGGGTTCATATGCCTGTGTGATGATTTCTGAATTCTGATACCGTT 255
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   114 TTGGACAAAATGCCCAAAGTCACCCCCACTACGGCTGATGTTTCCGGACCCGTTAAACTCCGTC 173
Qy   256 CACAAGGTGTTTGCTCAAGTTGCTAAGGATTTTTGTTAAGTTGCCATTGCACTTGGTTAAC 315
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   174 AACGAATGACCCTCTCCGATTATTATCTCCACGCACGGCCAAGAT---CGACAACCTGGTCACC 230
Qy   316 ACAGGTGGTTACTGTGAAAAACTTCCCATGTGAAAGATTACCCAGCACAAGAACGCTGAGAAG 375
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   231 TCCGCCGCGGATTCACGGAAAACTTCGATGCCATCCTCCTACCCTCACGGACCGTCTCGCAAAAG 290
Qy   376 ATGGTCAAGGTTAACTTGTTGTTGGGTTCTTTGTATGTTTCTCAAGGCCTTTGCTAAGCCATTG 435
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   291 CTTTGGGGCGTTAAATGTGATGGAACATACCTTTCGGCCACCGGTGTCGCGCAAAGC---AC 347
Qy   436 ATCAAAGAAGGTATCAAGGGTGCTTCGTGTTGTTTGATTGGTTCTATGTCTCGGTGGCCATT 495
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   348 CTCATCGGAGCGCCAAGGTTCCGGCAGCCATTGTCATCGATTGGTGCGATGTCTGGTGCTATC 407
Qy   496 GTCAACGATCCTCAAAAACCAAGTTGTCTACAACATGTCCAAGGCTGGGTGTTATCCATTTG 555
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   408 GACAACGTGCCGCAGCCCAGGGTCCTTACAACGCCGCCAAGGCCGCGCTGTTCGTCAACTT 467
Qy   556 GCTAAGACTTTGGCTTGTGTGAAGGGGCTAAGTACAACATCAGAGTTAATTCTTTAAACCA 615
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   468 GCCGCGGTACTTCGCGCGGGCGAATGGGCCCGGTCACGACATCCGGGTGAGCTGCATGCCCT 527
Qy   616 GGTTACATCTACGGTGTCCTTTGACCAAGAAATGTTATCAATGTGAACGGAAGAATTGTACAAC 675
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   528 CGATACATGCTTACTCGCGTCACCGACCCGAACAGATTTTGGATAGAACCCCGATTGCGGAC 587
Qy   676 AGATGGATCTCTGGTATCCCACACAACAAAGAATGTCCGAACCAAAAGGAATACAATGGTGCT 735
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   588 AAGTGGATCTCGCGTCATCCCCACCGGCCGAAGATGGGTACTCCGGAGGACCTGGATCGGTCCC 647
Qy   736 GTTTTGTACTTGCTTTCTGA 755
         ||||||||||||||||||
Db   588 AAGTGGATCTCGCGTCATCCCCACCGGCCGAAGATGGGTACTCCGGAGGACCTGGATCGGTCCC 647
Qy   648 GTTACCTTTCTGCTCAGNGA 667
```

RESULT 4

ID ADS61049 standard; cDNA; 637 BP.
XX
AC ADS61049;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #13036.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
PN US20032336754-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to
PT provide for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 36723; 122pp; English.
XX

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 637 BP; 149 A; 177 C; 169 G; 142 T; 0 U; 0 Other;

Query Match 10.0%; Score 81.8; DB 13; Length 637;
Best Local Similarity 55.0%; Pred. No. 3.1e-13;

Matches 183; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

Qy 304 CACTTGGTTAACACAGCTGGTTACTGTGAAAACTTCCCATGTGAAGATTACCCAGCCAAG 363
    ||||||| ||| || |||||||| ||| |||| ||| || ||| ||||| ||| |||
Db 70 CATCTCGTCACGCGTCTGCGCGGAGTTTACCGAAAAACTTCGACGCCATCAGCTACCCATACGAC 129

Qy 364 AACGCTGAGAGAAGATGGTGAAAGGTTAACTTGTGTTGGTTTCGTTGTTCTCAAGCCCTTT 423
    |||| ||| |||| || ||| ||| |||||||||| |||||||| ||||||||||||||
Db 130 CGTATGCAGAAGCTCTGGGGGGGTGAACGTGGACGGCACATATCTGTTTGCTACGGGCGGTG 189

Qy 424 GCTAAGCCATTGATCAAGAGAGGTATCAAGGGTGCTTCGTTGTTTTGATTGGTTCTATG 483
    ||| ||| || ||||| ||| ||| ||||| ||| ||||| |||| ||||||| |||
Db 190 GCAAAGCACTTGAT---GGAGCGCAAGGCGCCCGGAGTATTGTTGATCGGTAGCATG 246

Qy 484 TCTGGTGCCATTGTCAACGAATCCTCAAAACCAAGCTTGTTCTACACACATGTCCAAGGCTGGT 543
    |||| ||| |||||||||| || | | |||| |||| ||||| ||| ||| || ||||| |||
Db 247 TCTGGTGCCATCGTCAACGTACCTCAACCACAACGCCCTCAACGCGCCCAGCTCAACGAGGCAGCC 306

Qy 544 GTTATCCATTTGGCTAAGACTTTGGCTTGTGATGAGGGGTAAGTACACAATCAGAGTTAAT 603
    ||| ||| |||| ||| ||||| ||| ||||| ||||||||| || |||| |||||||
Db 307 GTTCGCCACCTTGCTTCGTTCGTCTTTCGCCCGTCGAATGGGCAGGACACATGGGTATCAGGGTAAAC 366

Qy 604 TCTTTAAACCCAGGTTACATCTACGGTTCCTTTG 636
    |||| || |||| ||| |||| ||| |||||
Db 367 TGCATTAGCCCTGGATACAATGCTGACTGCTCTG 399

RESULT 5

ID ADS48561 standard; cDNA; 735 BP.
XX
AC ADS48561;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #3304.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
PN US20032336754-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 26991; 122pp; English.
XX

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to

CC   provide for expression of a polynucleotide encoding a polypeptide from a
CC   microbial source. The invention also relates to a transformed plant
CC   comprising the recombinant DNA construct and a method of producing a
CC   transformed plant having an improved property. The plant is a crop plant
CC   such as maize or soybean. The method of producing a transformed plant
CC   having an improved property comprises transforming a plant with the
CC   recombinant DNA construct and growing the transformed plant, where the
CC   polynucleotide or polypeptide is useful for improving plant properties.
CC   The recombinant DNA construct is useful for producing plants with
CC   improved plant properties, e.g. improved cold, heat or drought tolerance,
CC   tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC   increased resistance to plant disease, better growth rate by modification
CC   of the cell cycle pathway with plant growth regulators, increased rate of
CC   homologous recombination, modified seed oil or protein yield and/or
CC   content, improved yield by modification of carbohydrate, nitrogen or
CC   phosphorus use and/or uptake, by modification of photosynthesis or by
CC   providing improved plant growth and development under at least one stress
CC   condition, improved lignin production or improved galactomannan
CC   production. This sequence represents a bacterial polynucleotide used in
CC   the scope of the invention. Note: The sequence data for this patent did
CC   not form part of the printed specification but was obtained in electronic
CC   format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 735 BP; 136 A; 234 C; 183 G; 182 T; 0 U; 0 Other;

Query Match   9.9%; Score 80.6; DB 13; Length 735;
Best Local Similarity   61.8%; Pred. No. 7.4e-13;
Matches 128; Conservative   0; Mismatches   79; Indels   0; Gaps   0;

Qy   436 ATCAAAGAAGGTATCAAGGGTGCTTCTGTTGTTTGTTGGTTCTATGTCTGGTGCCATT 495
          |||||| | || | || || || |||| ||| | || || | || | ||||
Db   364 ATGGTGGGCGCTGGCAAAACCGGGCTCCATTATCCTTGTGTCCGATGGTCCACCATC 423

Qy   496 GTCAACGATCCCAAAACCAAGTTGTCTACACAACATGTCCAAGGCTGGGTGGTTATCCATTG 555
          |||||| | ||| | | | | || ||| ||||| | ||||||||| | || ||| |||
Db   424 GTCAACTACCCCTCAGGAGCAGTCGTTGCTACAACCGCCTCTAAGGCTGGTGTCATTCAGCTC 483

Qy   556 GCTAAGGACTTTGGGCTTGTGAATGGGCTAAGTACACAACATCAGAGTTAATTCTTTAAACCCA 615
          | || | ||| | || | | | ||| |||||||| || ||| ||| |||| ||| |||||
Db   484 GGCAAGTCTCTTGCTGCCGGAGTGGGCCAAGTACAACAATCCGGTGTTAACTGCATTCCCC 543

Qy   616 GGTTACATCTACGGTCCTTTTGACCAAG 642
          |||||||| |||| ||| || | ||
Db   544 GGTTACATGGATACCGCGCCCTGAATAAG 570

RESULT 6
AAH74598
ID   AAH74598 standard; DNA; 852 BP.
XX
AC   AAH74598;
XX
DT   15-OCT-2001 (first entry)
XX
DE   Synthetic nucleotide sequence of a NADPH-dependent carbonyl reductase.
XX
KW   Methionine gamma-lyase; mdeA gene; free folding energy; gene shuffling;
KW   directed evolution; molecular breeding;
KW   NADPH-dependent carbonyl reductase; ss.
XX
OS   Candida magnoliae.
XX
PN   WO200155342-A2.
XX
PD   02-AUG-2001.
XX
PF   31-JAN-2001; 2001WO-US003186.
XX
PR   31-JAN-2000; 2000US-00499921.
PR   08-DEC-2000; 2000US-00734237.
XX
PA   (BIOC-) BIOCATALYTICS INC.

PI   Rozzell DJ, Bui P, Hua L;
XX
DR   WPI; 2001-483235/52.
XX
PT   Designing synthetic nucleic acid sequences for improved amplification,
PT   expression in host cell, by comparing free energy of folding of a
PT   starting polynucleotide and a modified polynucleotide having a codon
PT   replacement.
XX
PS   Claim 9; Page 104; 117pp; English.
XX
CC   The present sequence encodes a modified NADPH-dependent carbonyl
CC   reductase. The polynucleotide was modified using using the method of the
CC   invention. The specification describes a method for designing synthetic
CC   polynucleotide. The method comprises providing a starting polynucleotide,
CC   determining the predicted free energy of folding per base of the
CC   polynucleotide, modifying the polynucleotide by replacing a codon with a
CC   different codon to provide a modified polynucleotide, determining free
CC   energy of folding per base of the modified polynucleotide, and comparing
CC   this with that of the original polynucleotide. The method is useful for
CC   developing nucleic acid sequences that enhance expression of the encoded
CC   protein in a heterologous host. The design and preparation of the
CC   synthetic genes are used in application of gene shuffling, directed
CC   evolution and molecular breeding methods. The method allows expression of
CC   genes from various organisms such as mammals, plants, yeast, fungi and
CC   bacteria in prokaryotic hosts, such as Escherichia coli and eukaryotic
CC   hosts at commercially viable levels, in particular methionine gamma-lyase
CC   yield such as methionine gamma-lyase from P. putida
XX
SQ   Sequence 852 BP; 196 A; 216 C; 205 G; 235 T; 0 U; 0 Other;

Query Match   9.1%; Score 74.6; DB 5; Length 852;
Best Local Similarity   56.0%; Pred. No. 4.4e-11;
Matches 209; Conservative   0; Mismatches 149; Indels   15; Gaps   3;

Qy   437 TCAAAGAAGGTATCAAGGGTGCTTCTGTTGTTCTTGTTTTCAATTGGTTCTAATGTCTGGTGCCATTG 496
          || || || || | || ||| || || | || | | ||| | || | ||| || ||| | |
Db   485 TCGAAAAGGAAGGTPAAGAAGGGCTCGGTTTTCACCGCGTTCTAATGTCTCGGTCACATCG 544

Qy   497 TCAACGATCCTCAAAAACCAAGTTGTCTACAACAATGTCCAAGGCTGGGTGTTATCCATTTGG 556
          | || | ||| | | | | || || || || || | || ||| || | | | |||
Db   545 TTAACGTACCGCAGTTTCAGGCTACTACCAACAAGCGTGCTAAAGCTGGTGTGTCACTTCG 604

Qy   557 CTAAGACTTTGGCTTGTGAATGGGCTAAGTACACAACATCAGAGTTAATTCTTTAAACCCAG 616
          |||| | | ||| |||| | || ||||||||| |||||| || |||| || || | ||||
Db   605 CTAAAATCTCTGGCGTGTAGAATTCGGCTCCGT---TCGCTCGTGTTAACTCTGTTTCTCCGG 661

Qy   617 GTTACATCTACGGTCCTTTTGACCAAGAATGTTATCAATGGTAACGAAGAGATTGTTACAACA 676
          |||||| | ||| | | ||| | || || | |||| || | ||| ||||| ||||||
Db   662 GCTACAATCAAC------ACCGAAAATCTCGACTTTGTACCGCAGGAAAACTCAGAACA 712

Qy   677 GATGGAATCTCTGGTATCCCACAACACAAAAGAAATGTCCGGACCAAAGAGAATACATTCGGTCGTG 736
          ||||| || |||||| ||| |||| ||| || || | |||| || ||||| ||| ||| | ||
Db   713 AATGGTGGTCTCTGGTACCGCTGGGCGTGGTGGCGAAACTGCTGAACTGGTTGGTGCTT 772

Qy   737 TTTTGTTACTTGCTTTTCTGAATCTGCTTCATACGACTACTGGGTGCCCAGTTACTGGTTG 796
          || | |||| |||||| |||| |||| ||||| | || || || | || ||| | ||
Db   773 ACCTGTTTCTGGCTTCTGA---CGCTGGGTTCTTACGGCTACACCGGCACTGACATCATCGTTTG 829

Qy   797 ATGGTGGTTTCAC 809
          |||||||| | ||
Db   830 ACGGTGGTTACAC 842

RESULT 7
AAF12302
ID   AAF12302 standard; cDNA; 1146 BP.
XX
AC   AAF12302;
XX
DT   13-MAR-2001 (first entry)
XX
DE   Aspergillus oryzae EST SEQ ID NO:4825.

XX
KW   Multiple gene expression; filamentous fungal cell; EST;
KW   expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW   Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW   culture condition; environmental stress; spore morphogenesis;
KW   metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS   Aspergillus oryzae.
XX
PN   WO2005876702-A2.
XX
PD   28-SEP-2000.
XX
PF   22-MAR-2000; 2000WO-US007781.
XX
PR   22-MAR-1999;   99US-0027362J.
XX
PS   Claim 88; Page 2039; 316lpp; English.
XX
PT   Monitoring differential expression of genes in filamentous fungal cells
PT   uses fluorescence-labeled nucleic acids isolated from the cells and a
PT   substrate of expressed sequence tags.
XX
DR   WPI; 2000-594572/56.
XX
PI   Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX
PA   (NOVO) NOVO NORDISK BIOTECH INC.
PA   (NOVO) NOVO NORDISK AS.
XX
CC   The present invention describes a method for monitoring differential
CC   expression of genes in a first filamentous fungal (FF) cell relative to
CC   expression of the same genes in one or more second filamentous fungal
CC   cells. The method uses fluorescence-labeled nucleic acids isolated from
CC   the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC   are used in the methods for monitoring differential expression of genes
CC   in a first filamentous fungal (FF) cell relative to expression of the
CC   same genes in one or more second filamentous fungal cells. Monitoring the
CC   global expression of genes from FF cells allows the production potential
CC   of the microorganisms to be improved. New genes may be discovered,
CC   possible functions of unknown open reading frames can be identified and
CC   gene copy number variation and stability can be monitored. The expression
CC   of genes can be used to study how FF cells adapt to changes in culture
CC   conditions, environmental stress, spore morphogenesis, recombination,
CC   metabolic or catabolic pathway engineering. Using ESTs provides several
CC   advantages over genomic or random cDNA clones including elimination of
CC   redundancy as one spot on an array equals one gene or open reading frame,
CC   and organisation of the microarrays based on function of the gene
CC   products to facilitate analysis of the results. AAF07478 to AAF11247
CC   represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents
CC   ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from
CC   Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from
CC   Trichoderma reesei, which are all specifically claimed in the present
CC   invention
XX
SQ   Sequence 1146 BP; 262 A; 313 C; 274 G; 297 T; 0 U; 0 Other;

Query Match   9.1%; Score 74.2; DB 3; Length 1146;
Best Local Similarity   50.6%; Pred. No. 6.6e-11;
Matches   206; Conservative   0; Mismatches 198; Indels 3; Gaps 1;

Qy   349 GATTACCCAGCCAAGAACGCTGAGAAGATGGTGAAGGTTAACTTGTGTGGGTTCTTTGTAT 408
              |||| ||||||| ||||| |||||||||||||||||| |||| ||||||| | |||
Db   567 GAGTATACGGCCCAGGACCCAACACACGATGTTCGAAGTCAAGTCACTGGTGTGTTCATG 626

Qy   409 GTTTCTCAAGCCCTTTGCTAAGCCATTGATCAAAGAAGGTATCAAGGGGTGCTTCGTTGGTT 468
              | ||||||| ||| || | ||||||| | ||||||| ||| || ||||| |||| ||
Db   627 ACTTCCAAGCCCGTTGCTAAGCAAAATGATTCGGTTCCG---CAATGGAGGTAGCATCGCA 683

Qy   469 TTGATTGGTTCTATGTCTGGTGCCATTGTCAAACCAAGTTGTTCTACAC 528
              |||| ||| ||| | ||| |||||| ||| |||||||||||||
Db   684 CTAATTGCGGACGATGAGTGGTACTATATTGCCAATCGGGCCGCTTACAAAT 743

CC   tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC   increased resistance to plant disease, better growth rate by modification
CC   of the cell cycle pathway with plant growth regulators, increased rate of
CC   homologous recombination, modified seed oil or protein yield and/or
CC   content, improved yield by modification of carbohydrate, nitrogen or
CC   phosphorus use and/or uptake, by modification of photosynthesis or by
CC   providing improved plant growth and development under at least one stress
CC   condition. Improved lignin production or improved galactomannan
CC   production. This sequence represents a bacterial polynucleotide used in
CC   the scope of the invention. Note: The sequence data for this patent did
CC   not form part of the printed specification but was obtained in electronic
CC   format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 783 BP; 193 A; 186 C; 173 G; 231 T; 0 U; 0 Other;

Query Match   8.3%; Score 67.6; DB 13; Length 783;
Best Local Similarity   51.2%; Pred. No. 4.7e-09;
Matches 213; Conservative   0; Mismatches 194; Indels   9; Gaps   2;

Qy   171 CAAATACGGTTACTGAAGAATTGAAGTTGAAAGAAGTTCCAAAGAATGAGGGTTCATATGCCTG 230
                ||| |||| || || ||| ||||||||| |||||| ||||||||||| ||| |||
Db   171 CGAAAAGGCTGCAAAGGAAATTGCCCAGCGTAATGGGTGTCAAGACTTATACATGCCAAGTG 230

Qy   231 TGATATTTCTGATTCTGATACGGTTCACAAGGGTGTTGCTCAAGTTGCTAAGGATTTTGG 290
                |||| | | | ||||||||||||| |||||||||||||| | |
Db   231 CGATGTCACTATTCCTAAAGAGGTAGAACATGCCTTTGCCTTTGCTTTGCTGAGATTCAAAAAGGTCTTTGA 290

Qy   291 TAAGTTGCCATTGCACTTGGTTAACACAGCTGGTTACTGTGAAAACTTCCCACTGTGAAGA 350
                   ||| ||| | ||| ||| || || || ||| | || |||| || ||
Db   291 TAGTATTGATATCGTTGTGCCCAACACGGTATTTGCACTGGCAAGTCCCCATT---GA 347

Qy   351 TTACCCAGCCAAGAACGCTGAGAAGATGGTGAAGGTTAACTTGTTGGGTTCTTGTATGT 410
                || ||||| || || ||| | || ||| || ||| || |||| |||| | || ||
Db   348 CATGACTTACCGAGAAGAATTTGCTAATGAAAATAAACGTCAACTTGCTGCTCGGTGTCTTCAACGT 407

Qy   411 TTCTCAAGCCTTTGCTTAAGCCATTGATCAAAGAAGAGGTATCAAGGGGTGCTTCTGTTGGTTTT 470
                | ||||||| ||||| | |||| | ||||| || ||| | ||||||| ||| || |
Db   408 CGCTCACAATGCTGGCCCCCATCTTCCAAAAGCAAGG------CCATGGCAGCCTTGTACC 461

Qy   471 GAATGGTTCTATGTCTCGGTGGCCATTGTCAAACGATCCTCAAAAACCAAGTTGCTCTACAACAT 530
                |||||||| | |||| ||| |||||| ||| |||| |||||| ||||| || | |||||
Db   462 TACTGCTTCCATGTCTCGGTGTTGTTTGTTAACGTTGCTTCTCAACAACAATGTGCTTACAACAC 521

Qy   531 GTCCAAGGCTGGTGGTTATCCATTGGCTAAGACTTTGGGCTTGGTTGAATGGGCTAAGT 586
                ||||||| ||| ||| ||||| ||| | ||| |||| |||| | ||| |||||
Db   522 TTCCAAGGCGCGGTGTCATTCAACTCATCAAGAGTTGGCAGTGCCGCAAAT 577

XX
PS   Claim 1; SEQ ID NO 17; 71pp; English.

XX
PT   Increasing seed and organ size of a plant by transforming the plant with
PT   a DNA construct comprising a promoter that functions in plants and
PT   selecting a desired plant from a population of transformed plants
PT   containing the DNA construct.

XX
PA   (MONS) MONSANTO TECHNOLOGY LLC.
XX
PI   Fernandes M, Xie Z, Dotson SB;
XX
DR   WPI; 2004-042481/04.
DR   P-PSDB; ADN62513.

XX
AC   ADN62512;
XX
DT   01-JUL-2004   (first entry)
XX
DE   A. thaliana gene SEQ ID NO:17.
XX
KW   de; gene; seed size; organ size; plant; transgenic.
XX
OS   Arabidopsis thaliana.

XX
FH   Key   Location/Qualifiers
FT   CDS   1..789
FT   /*tag=   a
XX

XX
SQ   Sequence 789 BP; 193 A; 186 C; 173 G; 231 T; 0 U; 0 Other;

XX
PN   WO2003096797-A2.
XX
PD   27-NOV-2003.
XX
PF   14-MAY-2003; 2003WO-US014989.
XX
PR   15-MAY-2002; 2002US-0381100P.

XX
AC   ADN62512;
ID   ADN62512 standard; DNA; 789 BP.

RESULT 9
ADN62512

The invention relates to a novel method for increasing seed and organ
size of a plant by: transforming the plant with a DNA construct
comprising a promoter that functions in plants, operably linked to a DNA
molecule that encodes a protein; and selecting a desired plant from a
population of transformed plants containing the DNA construct; where the
desired plant exhibits increased seed and organ size compared to a plant
of a same plant species not transformed to contain the DNA construct. The
method of the invention is useful for increasing seed and organ size of a
plant. The present sequence is used in the exemplification of the
invention.

Query Match   8.1%; Score 66.4; DB 12; Length 789;
Best Local Similarity   56.5%; Pred. No. 1.1e-08;
Matches 165; Conservative   0; Mismatches 121; Indels   6; Gaps   2;

Qy   519 TGTCTACACATGTCCAAGGCTGGTGTTATCCATTGGCTAAGACTTTGGGCTTGGTGAATG 578
                || ||| | ||| ||||| ||| ||||| || |||| |||| |||| | | |||
Db   471 TATCTATAGTTTAACGAAAGGAGCTTTGAATCAACTAGCAAAAAACTTTGGCCATGTGAAATG 530

Qy   579 GGCTAAGTACAACATCAGAGAGTTAATTCTTTAAACCCAGGTTACATCTACAGGGTCTTTTGAC 638
                ||| |||||||| || || ||| | | ||| | ||||||| | |||| | || | |||||
Db   531 GGCAAGAGATGGCATAAGAGCCAACTCTGTTGCTCCTAATTTAACCTACACTGCTATGGC 590

Qy   639 CAAGAATGTTATCAATGGGTAACGAAGAAGTTGTACAACAGAATGGATCTCTGGTATCCCACA 698
                |||||||| | || ||| |||| |||||||| ||||| || ||||| || ||| | |||
Db   591 TCAACCTTTTTCAAAG---ACGCCGATTACGGACGAAGAAGAGTTTGGTTAGTTAGAACTCACT 647

Qy   699 ACAAAGAAATGTCCGACCAAAGGAATAACAATGGTCGTCGTTTTGTTACTTGCTTCTGAATC 758
                ||| | | ||| || ||||| | ||| | ||| |||| |||| | ||| || | || ||
Db   648 TGGTGCGCTGGAGAGCCAAAATGAGG---TTTCATCACTGGTTGGGCTTTTCTGTGTCTACC 704

Qy   759 TGCTCGCTTCATACACTACTACTGGGTGCCAGCTTACTGGTTGATGGGTTGGTTTCACT 810
                |||| |||| | | |||||| ||| | | ||||| |||| ||| ||||| |||||||
Db   705 TGCACGTTCATTATATTACTACTGGGTCAGACCCAATTTGTCGTTGAAGGACGAGGTCTCACT 756

XX
AC   ADS48180;
XX
DT   02-DEC-2004   (first entry)

RESULT 10
ADS48180
ADS48180 standard; cDNA; 774 BP.

DE   Bacterial polynucleotide #2923.
XX
KW   Recombinant DNA construct; transformed plant; improved plant property;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW   pathogen tolerance; pest tolerance; plant disease resistance;
KW   cell cycle pathway modification; plant growth regulator;
KW   homologous recombination; seed oil yield; protein yield; carbohydrate;
KW   nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW   bacterial polynucleotide; gene; ss.
XX
OS   Bacteria.

PN US20030233675-A1
XX
PD 18-DEC-2003.
XX
PF
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PS Claim 1; SEQ ID NO 26610; 122pp; English.
XX

CC The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The method of producing a transformed plant such as maize or soybean, comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the recombinant DNA construct and polypeptide or polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, improved resistance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, improved seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polynucleotide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.
XX

SQ Sequence 774 BP; 226 A; 160 C; 229 G; 159 T; 0 U; 0 Other;

Query Match   7.8%; Score 63.6; DB 13; Length 774;
Best Local Similarity   61.4%;
Matches 102; Conservative   0; Mismatches   64; Indels   0; Gaps   0;

Qy   472 ATTGGTTCTATGTCTGGTGCCATTGTCAAGCGAACCCTCAAAACCAAGTTGTCTCAACATG 531
          || ||| | |||||| || | ||||| || |||| | ||||| ||| ||| |
Db   436 ATCGGCGTCCATGTCCGGACACATCGTGAACAAAACCTCAGAAGCAGACAGCTTACAACGCT 495

Qy   532 TCCAAGGCTGGTGTTATCCATTGGCCTTGTGTGAATGGGCTAAGTACAAC 591
          || |||| ||| |||| | ||| | | | || | ||| |||||| | |
Db   496 TCGAAAGCGGGGTGTGATCCATCTCACCAGATCTTGGCGCCGCCGTACCGA 555

Qy   592 ATCAGAGGTTAATTCTTTAAACCCAGGTTACATCTTACGGGTTCCTTTTGA 637
          ||||| ||||| | ||| ||||| ||||||| ||||||||||| | |
Db   556 ATCAGGGGTGAACAGCAATAAGCCCGGATACACTTACGGGTTCCTTCTCA 601

DT 13-MAR-2001 (first entry)
XX
DE Aspergillus niger EST SEQ ID NO:3866.
XX
KW Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Aspergillus niger.
XX
PN WO200056762-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US007781.
XX
PR 22-MAR-1999; 99US-00273623.
XX
DR WPI; 2000-594572/56.
XX
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO NORDISK AS.
XX
PT Monitoring differential expression of genes in filamentous fungal cells uses fluorescence-labeled nucleic acids isolated from the cells and a substrate of expressed sequence tags.
XX
PS Claim 87; Page 1728; 3161pp; English.
XX

CC The present invention describes a method for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. The method uses fluorescence-labeled nucleic acids isolated from the FF cells and a substrate of expressed sequence tags (EST). The ESTs are used in the methods for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. Monitoring the global expression of genes from FF cells allows the production potential of the microorganisms to be improved. New genes may be discovered, possible functions of unknown open reading frames can be identified and gene copy number variation and stability can be monitored. The expression of genes can be used to study how FF cells adapt to changes in culture conditions, environmental stress, spore morphogenesis, recombination, metabolic or catabolic pathway engineering. Using ESTs provides several advantages over genomic or random cDNA clones including elimination of redundancy as one spot on an array equals one gene or open reading frame, and organisation of the microarrays based on function of the gene products to facilitate analysis of the results. AAF07478 to AAF11247 represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are all specifically claimed in the present invention
XX

SQ Sequence 636 BP; 169 A; 155 C; 161 G; 149 T; 0 U; 2 Other;

Query Match   7.4%; Score 60.6; DB 3; Length 636;
Best Local Similarity   51.3%; Pred. No. 4.7e-07;
Matches 194; Conservative   0; Mismatches 175; Indels   9; Gaps   2;

Qy   211 AAGATGGGTTCATATGCCTGTGATATTTCTGATTCTGATACCGTTCACAAGGTGTTTGCT 270
          ||| | ||| |||| | || ||| | | ||| ||| | | || | || | | ||
Db   200 AAGTCAGCTGCCTATTAAAGCTGAGTTGTCGGAAATCAGGAGGATATTTGAGAGGGTAACA 259

Qy   271 CAAGTTGCTAAGGATTTTGGTAAGTTGCCCATGCCACTTGGTTAACACAGCTTGGTTACTGT 330
          || |||||| ||| | ||||| ||| || | | | | || | | || |||| |||
Db   260 CAGATTGCTTCAGACTTTGGCAAGCTTGATATTATATCGTCGTGGAAC---TCGGGGGTCACT 316

Qy 331 GAAAACTTCCCATGTGAAGATTACCCAGCCAAGAACGCTGAGAGAGATGGTGAAGAGGTTAAC 390
        ||||| || ||| || || || || | || || || | | || || ||
Db 317 TCCAATATCGCGGCAGAAGACTACACCACCAGAACAATGGCGGTGACATCATGAAGGTCAAT 376

Qy 391 TTGTTGGGTTCTTCGTTTGTATGTTTCTCAAGCCTTTGCTTAAGCCATTGAATCAAAGAGGTATC 450
        | | |||| | || | || || || ||| | ||| || || || || | ||
Db 377 CTAGATGCGCGCATTTTATACGCGCAACGCCGGCTTCTCGGATCTTCAAGCAACACAGG---- 432

Qy 451 AAGGGTGCTTCTGTTGTTTTGATTGTTGGTCCCATTGTCAACGATCCTCAA 510
        | | ||| | || | | | | ||| || ||| | || ||
Db 433 --ACATGGAAATGTAATCTTCACAGCCTTGCAGTGCAACATTGGTGAAAGACTTTGGCT 490

Qy 511 AACCAAGTTGTCTCACAACATGTCCAAGGCGTGGTGTTATCCATTTGGGCTAAGACTTTGGCT 570
        || |||| | | | |||| || || || ||| || |||| || | ||| || |||||
Db 491 AAGCAAGCTGCATACAATGCCTCCAAAGCTGTGGGGGTCGTGCAAAATGGCAAAAATGTCTGTTCT 550

Qy 571 TGTGAATGGGCTAAGTAC 588
        ||||||||| ||||
Db 551 GTTGAATGGGTTGANTTC 568

RESULT 12
AAL57425
ID   AAL57425 standard; DNA; 822 BP.
XX
AC   AAL57425;
XX
DT   09-OCT-2003 (first entry)
XX
DE   Datura stramonium tropinone reductase-1 gene.
XX
KW   Optically active alcohol; assymetric ketone reduction; jimsonweed;
KW   tropinone reductase-1; reduced co-enzyme; (R)-3 quinuclidinol;
KW   arteriosclerosis; co-expression; common thornapple; gene; ds.
XX
OS   Datura stramonium.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..822
FT   /*tag= a
FT   /product= "Datura stramonium tropinone reductase-1"
XX
PN   EP1318200-A2.
XX
PD   11-JUN-2003.
XX
PF   06-DEC-2002; 2002EP-00027311.
XX
PR   07-DEC-2001; 2001JP-00375041.
PR   27-MAY-2002; 2002JP-00152955.
PR   31-MAY-2002; 2002US-0385434P.
XX
PA   (DAIL) DAICEL CHEM IND LTD.
XX
PI   Yamamoto H, Ueda M, Ritsuzui P, Hamatani T;
XX
PT   Preparation of optically active alcohols, particularly quinuclidinols,
PT   via asymmetric reduction of a ketone, e.g. 3 quinuclidinone, using
PT   tropinone reductase-1.
XX
PS   Claim 9; Page 20-21; 30pp; English.
XX
CC   This invention relates to a novel method for the preparation of an
CC   optically active alcohol by assymetric reduction of a ketone with an
CC   enzymatic material having tropinone reductase-1 activity, in the presence
CC   of a reduced co-enzyme. Optically active alcohols are useful in the
CC   preparation of optically active compounds, for example (R)-3
CC   quinuclidinol is an intermediate in the preparation of therapeutic
CC   agents. These can be used, for example, for the treatment of
CC   arteriosclerosis. Co-expression of the 2 enzymes, as in the method of the

CC   invention, enables more efficient production of optically active
CC   alcohols. The present sequence is the Datura stramonium (jimsonweed;
CC   common thornapple) tropinone reductase-1 gene sequence. The protein
CC   encoded by this gene was utilised in the method of the invention
XX
SQ   Sequence 822 BP; 255 A; 140 C; 183 G; 244 T; 0 U; 0 Other;

Query Match   7.3%; Score 59.8; DB 9; Length 822;
Best Local Similarity   52.5%; Pred. No. 9e-07;
Matches 158; Conservative   0; Mismatches 137; Indels   6; Gaps   1;

Qy 519 TGTCTCAACACATGTCCAAGGCGTGGTGTTATCCATTTGGGCTAAGACTTTGGCTTGTGAATG 578
        ||| || | ||||||| || ||||| |||||| ||||| |||||||| ||||||
Db 507 TCTTTACTCAGCTTCCAAAGCTGCAATAAATCAAATGACAAAGAGTTTGGCTGTGAAATG 566

Qy 579 GGCTAAGTACAAACATCAGAGGTTAATTCTTTAAACCCAGGTTACATCTACGGTGTCCTTTGAC 638
        |||| | |||| ||| ||| ||| || | || | || || | || | | ||| ||
Db 567 GGCTAAGACAACAATTCGGGTCAATTCAGTTGCTCGGGACTGATTTTAACCCCACTGGT 626

Qy 639 CAAGAATGTTATC------AATGGTAACGAAGAATTGTACAACAGATGGATCTCTGGTTAT 692
        |||||||| ||| ||||| ||| ||||| ||| | || | || | || ||| |
Db 627 TGAAACTGCAATTAAGAAAATCCTCATCAAAAGAAATAGCAAATTTTATTGTCAA 686

Qy 693 CCCACAACAAAGAATGTCCGAACCAAAAGGAATACAATCATTGGTGCTGTTTGTACTTGCTTC 752
        ||| |||| ||||||| || ||| ||||| | ||| || | | ||| || || || ||
Db 687 GACTCCTATGGCGCCGGCAAAGCCCCAGAAAGTTCTGCACTAATAGCTTTTCTTG 746

Qy 753 TGAATCTGCTGCTTCATCACTACTACTGGTGCCAGCTTACTCGGTTGATGGGTGGTTTCACTTC 812
        ||| | ||| ||||| | | | || | ||||| ||| | | |||| | ||| ||||| ||
Db 747 CTTCCCTGCTCCTTCATATATTAATGGGCTGAACTCATATATGGGCTGAGTTCACAGC 806

Qy 813 T 813
        |
Db 807 T 807

RESULT 13
AAZ46762
ID   AAZ46762 standard; DNA; 774 BP.
XX
AC   AAZ46762;
XX
DT   31-MAR-2000 (first entry)
XX
DE   Bacillus D-arabinitol dehydrogenase coding sequence.
XX
KW   D-arabinitol dehydrogenase; clinical diagnosis; mycosis; ds.
XX
OS   Bacillus sp.
XX
FH   Key   Location/Qualifiers
FT   CDS   1
FT   /*tag= a
FT   /product= "D-arabinitol dehydrogenase"
FT   /note= "the stop codon is not indicated"
XX
PN   JP11332569-A.
XX
PD   07-DEC-1999.
XX
PF   26-MAY-1998; 98JP-00143637.
XX
PR   26-MAY-1998; 98JP-00143637.
XX
PA   (IKED-) IKEDA SHOKKEN KK.
PA   (NIPK) NIPPON KAYAKU KK.
XX
PT   Arabinitol dehydrogenase gene encoding D-arabinitol dehydrogenase -
PT   useful as a clinical diagnosing agent for mycosis.
XX
DR   WPI; 2000-091353/08.
DR   P-PSDB; AAY56815.
XX

Left column

```
PS   Claim 1; Page 10-11; 14pp; Japanese.
PT   Arabinitol dehydrogenase.
XX
CC   The invention relates to gene encoding D-arabinitol dehydrogenase,
CC   isolated from Bacillus sp. IKD-5A868 strain. The protein can be expressed
CC   by standard recombinant methodologies. D-arabinitol dehydrogenase is used
CC   as a clinical diagnosing agent for mycosis. The present sequence
CC   represents the coding sequence of the D-arabinitol dehydrogenase
XX
SQ   Sequence 774 BP; 155 A; 214 C; 271 G; 134 T; 0 U; 0 Other;

     Query Match       7.2%;  Score 59;  DB 3;  Length 774;
     Best Local Similarity   50.4%;  Pred. No. 1.5e-06;
     Matches  171;  Conservative   0;  Mismatches  165;  Indels   3;  Gaps   1;

Qy   478 TCTATGTCTCGGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTCTACAACATGTCCAAG 537
             ||||| |||| | |||  ||||| ||    |||  ||    |||||||| || || |||||
Db   439 TCCATGTCCGGGCCTGATCGTCAATACACGGCGCAGCGCGCCGCCTACAACATGTCTCCAAG 498

Qy   538 GCTGGTGTTATCCATTGGCTAAGACTTTGGCTTGGAATGGGCTAAGTACAACATCAGA 597
             ||  | |||  || ||| | |  |||  || ||||||||| || ||  | |||||||
Db   499 GCCGGGGGTCATTATGCTGACCAAGAGCCTGGCATCCGAATGGGCGCCCGCACGGGTGCGC 558

Qy   598 GTTAATTCTTTAAACCCAGGTTACATCTACCGGTCCTTTGACCAAGAATGTTATCAAATGGT 657
             || || |  | |  | | ||| ||| | | | ||  || || ||||| | | |  |||
Db   559 GTCAACACGATTGCGCCCCGGCTACATGAAGACGAAGCTGACGGAACCGTATTTCGCGCGCC 618

Qy   658 AACGAAGAATTGTACACAGATGGATCTCTGGTATCCCACACAAAAGAAGTTCCGAACCA 717
             | ||| ||| | |  | |||| |  ||  || |||  ||  ||||| | ||||||||
Db   619 GGAGGCGGAGATGAGATGGCTGGCTATGAGCCGTTCC- 677

Qy   718 AAGGAAATACAATTGGTGCGTGTTTTGTACTTGCTTTCTGAATCTCGTCGTCATACACTACT 777
             |||||||||| || ||| |   |  ||| || ||  ||||| | ||  || || ||  | |
Db   678 --GCACGAGCTGGGCGCCAATTGCGGCGTTCACCTTTGCTTCCTTCGGACGGCTTCC- 735

Qy   778 GGTGCCCAGCTTACTTGGTTGATGGTGGTTTCACTTCTTGG 816
             || || |  ||| | |||   | || || |||||  ||||
Db   736 GGCGCGCGTGTTCACGATTGACGGGGCGGGGCTATACCGATTTCGG 774


RESULT 14
AA246763
ID   AA246763 standard; DNA; 2255 BP.
XX
AC   AA246763;
XX
DT   31-MAR-2000 (first entry)
XX
DE   Bacillus D-arabinitol dehydrogenase encoding genomic DNA.
XX
KW   D-arabinitol dehydrogenase; clinical diagnosis; mycosis; ds.
XX
OS   Bacillus sp.
XX
FH   Key            Location/Qualifiers
FT   CDS            366..1142
FT                  /*tag=   a
FT                  /product= "D-arabinitol dehydrogenase"
FT                  /note= "the coding sequence is also given in AA246762"
XX
PN   JP11332569-A.
XX
PD   07-DEC-1999.
XX
PF   26-MAY-1998;  98JP-00143637.
XX
PR   26-MAY-1998;  98JP-00143637.
XX
PA   (IKED-) IKEDA SHOKKEN KK.
PA   (NIPK ) NIPPON KAYAKU KK.
XX
DR   WPI; 2000-091353/08.
DR   N-PSDB; AAY56815.
```

Right column

```
PT   Arabinitol dehydrogenase gene encoding D-arabinitol dehydrogenase -
PT   useful as a clinical diagnosing agent for mycosis.
XX
PS   Claim 6; Page 12; 14pp; Japanese.
XX
CC   The invention relates to gene encoding D-arabinitol dehydrogenase,
CC   isolated from Bacillus sp. IKD-5A868 strain. The protein can be expressed
CC   by standard recombinant methodologies. D-arabinitol dehydrogenase is used
CC   as a clinical diagnosing agent for mycosis. The present sequence
CC   represents the DNA encoding the D-arabinitol dehydrogenase
XX
SQ   Sequence 2255 BP; 418 A; 623 C; 728 G; 486 T; 0 U; 0 Other;

     Query Match       7.2%;  Score 59;  DB 3;  Length 2255;
     Best Local Similarity   50.4%;  Pred. No. 2.4e-06;
     Matches  171;  Conservative   0;  Mismatches  165;  Indels   3;  Gaps   1;

Qy   478 TCTATGTCTCGGTGCCATTGTCAACGATCCTCAAAAACCAAGTTGTCTACAACATGTCCAAG 537
             ||||| |||| | |||  ||||| ||    |||  ||    |||||||| || || |||||
Db   804 TCCATGTCCGGGCCTGATCGTCAATACACGGCGCAGCGCGCCGCCTACAACATGTCTCCAAG 863

Qy   538 GCTGGTGTTATCCATTGGCTAAGACTTTGGCTTGGAATGGGCTAAGTACAACATCAGA 597
             ||  | |||  || ||| | |  |||  || ||||||||| || ||  | |||||||
Db   864 GCCGGGGGTCATTATGCTGACCAAGAGCCTGGCATCCGAATGGGCGCCCGCACGGGTGCGC 923

Qy   598 GTTAATTCTTTAAACCCAGGTTACATCTACCGGTCCTTTGACCAAGAATGTTATCAAATGGT 657
             || || |  | |  | | ||| ||| | | | ||  || || ||||| | | |  |||
Db   924 GTCAACACGATTGCGCCCCGGCTACATGAAGACGAAGCTGACGGAACCGTATTTCGCGCGCC 983

Qy   658 AACGAAGAATTGTACACAGATGGATCTCTGGTATCCCACACAAAAGAAGTTCCGAACCA 717
             | ||| ||| | |  | |||| |  ||  || |||  ||  ||||| | ||||||||
Db   984 GGAGGCGGAGATGAGATGGCTGGCTATGAGCCGTTCC- 1042

Qy   718 AAGGAAATACAATTGGTGCGTGTTTTGTACTTGCTTTCTGAATCTCGTCGTCATACACTACT 777
             |||||||||| || ||| |   |  ||| || ||  ||||| | ||  || || ||  | |
Db  1043 --GCACGAGCTGGGCGCCAATTGCGGCGTTCACCTTTGCTTCCTTCGGACGGCTTCC- 1100

Qy   778 GGTGCCCAGCTTACTTGGTTGATGGTGGTTTCACTTCTTGG 816
             || || |  ||| | |||   | || || |||||  ||||
Db  1101 GGCGCGCGTGTTCACGATTGACGGGGCGGGGCTATACCGATTTCGG 1139


RESULT 15
AAC42189
ID   AAC42189 standard; DNA; 798 BP.
XX
AC   AAC42189;
XX
DT   17-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 34619.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000;  2000EP-00301439.
XX
PR   25-FEB-1999;  99US-0121825P.
PR   05-MAR-1999;  99US-0123180P.
PR   09-MAR-1999;  99US-0123548P.
PR   23-MAR-1999;  99US-0125788P.
PR   25-MAR-1999;  99US-0126264P.
PR   25-MAR-1999;  99US-0126264P.
PR   29-MAR-1999;  99US-0126785P.
PR   01-APR-1999;  99US-0127462P.
PR   01-APR-1999;  99US-0127462P.
PR   06-APR-1999;  99US-0128234P.
PR   06-APR-1999;  99US-0128234P.
PR   08-APR-1999;  99US-0128714P.
```

PR 16-APR-1999; 99US-0129845P.	PR 19-JUL-1999; 99US-0144334P.
PR 19-APR-1999; 99US-0130077P.	PR 19-JUL-1999; 99US-0144335P.
PR 21-APR-1999; 99US-0130449P.	PR 20-JUL-1999; 99US-0144352P.
PR 23-APR-1999; 99US-0130510P.	PR 20-JUL-1999; 99US-0144632P.
PR 28-APR-1999; 99US-0130891P.	PR 20-JUL-1999; 99US-0144884P.
PR 30-APR-1999; 99US-0131449P.	PR 21-JUL-1999; 99US-0144814P.
PR 30-APR-1999; 99US-0132048P.	PR 21-JUL-1999; 99US-0145086P.
PR 04-MAY-1999; 99US-0132407P.	PR 21-JUL-1999; 99US-0145088P.
PR 04-MAY-1999; 99US-0132484P.	PR 22-JUL-1999; 99US-0145085P.
PR 05-MAY-1999; 99US-0132485P.	PR 22-JUL-1999; 99US-0145087P.
PR 06-MAY-1999; 99US-0132486P.	PR 22-JUL-1999; 99US-0145089P.
PR 07-MAY-1999; 99US-0132487P.	PR 22-JUL-1999; 99US-0145192P.
PR 07-MAY-1999; 99US-0132863P.	PR 23-JUL-1999; 99US-0145145P.
PR 11-MAY-1999; 99US-0134256P.	PR 23-JUL-1999; 99US-0145218P.
PR 14-MAY-1999; 99US-0134218P.	PR 23-JUL-1999; 99US-0145224P.
PR 14-MAY-1999; 99US-0134219P.	PR 26-JUL-1999; 99US-0145276P.
PR 14-MAY-1999; 99US-0134221P.	PR 27-JUL-1999; 99US-0145913P.
PR 18-MAY-1999; 99US-0134370P.	PR 27-JUL-1999; 99US-0145918P.
PR 19-MAY-1999; 99US-0134768P.	PR 28-JUL-1999; 99US-0145919P.
PR 20-MAY-1999; 99US-0134941P.	PR 04-AUG-1999; 99US-0146386P.
PR 21-MAY-1999; 99US-0135124P.	PR 02-AUG-1999; 99US-0146388P.
PR 24-MAY-1999; 99US-0135353P.	PR 02-AUG-1999; 99US-0146389P.
PR 25-MAY-1999; 99US-0135629P.	PR 03-AUG-1999; 99US-0147038P.
PR 27-MAY-1999; 99US-0136021P.	PR 04-AUG-1999; 99US-0147204P.
PR 28-MAY-1999; 99US-0136392P.	PR 04-AUG-1999; 99US-0147302P.
PR 01-JUN-1999; 99US-0136782P.	PR 05-AUG-1999; 99US-0147192P.
PR 03-JUN-1999; 99US-0137222P.	PR 05-AUG-1999; 99US-0147260P.
PR 04-JUN-1999; 99US-0137528P.	PR 06-AUG-1999; 99US-0147303P.
PR 07-JUN-1999; 99US-0137502P.	PR 06-AUG-1999; 99US-0147416P.
PR 08-JUN-1999; 99US-0137724P.	PR 09-AUG-1999; 99US-0147493P.
PR 10-JUN-1999; 99US-0138094P.	PR 09-AUG-1999; 99US-0147935P.
PR 14-JUN-1999; 99US-0138540P.	PR 10-AUG-1999; 99US-0148171P.
PR 16-JUN-1999; 99US-0138847P.	PR 11-AUG-1999; 99US-0148319P.
PR 17-JUN-1999; 99US-0139119P.	PR 12-AUG-1999; 99US-0148341P.
PR 18-JUN-1999; 99US-0139453P.	PR 13-AUG-1999; 99US-0148565P.
PR 18-JUN-1999; 99US-0139492P.	PR 13-AUG-1999; 99US-0148684P.
PR 18-JUN-1999; 99US-0139452P.	PR 16-AUG-1999; 99US-0149368P.
PR 18-JUN-1999; 99US-0139454P.	PR 17-AUG-1999; 99US-0149175P.
PR 18-JUN-1999; 99US-0139455P.	PR 18-AUG-1999; 99US-0149426P.
PR 18-JUN-1999; 99US-0139456P.	PR 20-AUG-1999; 99US-0149722P.
PR 18-JUN-1999; 99US-0139457P.	PR 20-AUG-1999; 99US-0149723P.
PR 18-JUN-1999; 99US-0139458P.	PR 20-AUG-1999; 99US-0149929P.
PR 18-JUN-1999; 99US-0139459P.	PR 23-AUG-1999; 99US-0149902P.
PR 18-JUN-1999; 99US-0139460P.	PR 23-AUG-1999; 99US-0149930P.
PR 18-JUN-1999; 99US-0139461P.	PR 25-AUG-1999; 99US-0150566P.
PR 18-JUN-1999; 99US-0139462P.	PR 26-AUG-1999; 99US-0150884P.
PR 18-JUN-1999; 99US-0139463P.	PR 27-AUG-1999; 99US-0151065P.
PR 21-JUN-1999; 99US-0139750P.	PR 27-AUG-1999; 99US-0151066P.
PR 22-JUN-1999; 99US-0139763P.	PR 30-AUG-1999; 99US-0151080P.
PR 23-JUN-1999; 99US-0139817P.	PR 31-AUG-1999; 99US-0151303P.
PR 24-JUN-1999; 99US-0139899P.	PR 01-SEP-1999; 99US-0151438P.
PR 28-JUN-1999; 99US-0140353P.	PR 07-SEP-1999; 99US-0151930P.
PR 29-JUN-1999; 99US-0140354P.	PR 10-SEP-1999; 99US-0152363P.
PR 30-JUN-1999; 99US-0140695P.	PR 13-SEP-1999; 99US-0153070P.
PR 01-JUL-1999; 99US-0140823P.	PR 15-SEP-1999; 99US-0153758P.
PR 02-JUL-1999; 99US-0140991P.	PR 16-SEP-1999; 99US-0154018P.
PR 06-JUL-1999; 99US-0141287P.	PR 20-SEP-1999; 99US-0154039P.
PR 08-JUL-1999; 99US-0141842P.	PR 22-SEP-1999; 99US-0154779P.
PR 09-JUL-1999; 99US-0142154P.	PR 23-SEP-1999; 99US-0155139P.
PR 12-JUL-1999; 99US-0142055P.	PR 24-SEP-1999; 99US-0154486P.
PR 13-JUL-1999; 99US-0142390P.	PR 28-SEP-1999; 99US-0155659P.
PR 14-JUL-1999; 99US-0142803P.	PR 29-SEP-1999; 99US-0154658P.
PR 15-JUL-1999; 99US-0142920P.	PR 04-OCT-1999; 99US-0156596P.
PR 16-JUL-1999; 99US-0142977P.	PR 05-OCT-1999; 99US-0157117P.
PR 19-JUL-1999; 99US-0143542P.	PR 06-OCT-1999; 99US-0157553P.
PR 19-JUL-1999; 99US-0143624P.	PR 07-OCT-1999; 99US-0157865P.
PR 19-JUL-1999; 99US-0144005P.	PR 08-OCT-1999; 99US-0158029P.
PR 19-JUL-1999; 99US-0144086P.	PR 12-OCT-1999; 99US-0158322P.
PR 19-JUL-1999; 99US-0144325P.	PR 13-OCT-1999; 99US-0158369P.
PR 19-JUL-1999; 99US-0144331P.	PR 13-OCT-1999; 99US-0159293P.
PR 19-JUL-1999; 99US-0144332P.	PR 13-OCT-1999; 99US-0159294P.
PR 19-JUL-1999; 99US-0144333P.	PR 13-OCT-1999; 99US-0159295P.

```
PR   14-OCT-1999;    99US-0159329P.
PR   14-OCT-1999;    99US-0159330P.
PR   14-OCT-1999;    99US-0159331P.
PR   14-OCT-1999;    99US-0159637P.
PR   14-OCT-1999;    99US-0159638P.
PR   18-OCT-1999;    99US-0159584P.
PR   21-OCT-1999;    99US-0160741P.
PR   21-OCT-1999;    99US-0160767P.
PR   21-OCT-1999;    99US-0160768P.
PR   21-OCT-1999;    99US-0160770P.
PR   21-OCT-1999;    99US-0160814P.
PR   21-OCT-1999;    99US-0160815P.
PR   22-OCT-1999;    99US-0160900P.
PR   22-OCT-1999;    99US-0160980P.
PR   22-OCT-1999;    99US-0160981P.
PR   22-OCT-1999;    99US-0160989P.
PR   25-OCT-1999;    99US-0161404P.
PR   25-OCT-1999;    99US-0161405P.
PR   25-OCT-1999;    99US-0161406P.
PR   25-OCT-1999;    99US-0161359P.
PR   26-OCT-1999;    99US-0161360P.
PR   26-OCT-1999;    99US-0161361P.
PR   26-OCT-1999;    99US-0161920P.
PR   28-OCT-1999;    99US-0161922P.
PR   28-OCT-1999;    99US-0161993P.
PR   29-OCT-1999;    99US-0162142P.

Query Match        6.9%;    Score  56.4;   DB 3;   Length 798;
Best Local Similarity   54.5%;   Pred. No. 8.8e-06;
Matches  158;  Conservative   0;  Mismatches  126;   Indels    6;   Gaps   2;

Qy   521  TCTACAACATGTCCAAGCCTGGGTGTTATCCATTTGGCTAAGACTTTGGCTTGTGAATGGG  580
          |       ||||  ||| |  |   ||||  | |     |||  | | ||| ||| ||||||
Db   482  TTTATAGTCTAACTAAAAGGAGACTTTGAACCAGTTAACTAGAAAATTTAGCATGTGAATGGG  541

Qy   581  CTAAGTACAACATCAGAGTTAATTCTTTAAACCCAGGTTACATCTACCGGTCCTTTGACCA  640
          | ||||| |||| |||| | ||  | |    | ||||||||||||  ||||||||||||||
Db   542  CAAAAGACGGGCATAAGACCCAATGCTGTTGCACCTAATGTTGGTCAAGACTCCTTTGTCTC  601

Qy   641  AGAATGTTATCAATGGTAACGAAGAAATTGTACAACAGAGATGGATCTCTGGTATCCCACAAC  700
          |   ||||| |  | | | | || | |  | |  | || |  ||||| ||| | || || ||
Db   602  AATCTTATCTCGAGG---ACGTCGGTTTCAAGGAGGCATTGTTCAGTAGAACTCCACTTG  658

Qy   701  AAAGAAATGTCCGAACCAAAAGGAATACATTGGGTGCTGTTTTGTACTTGCTTTTCTGAAATCTG  760
          |        ||||| |||         |   |   |||| | |||| ||| ||||||||| |||
Db   659  GTCGCGCTGGAGAGCCGAAATGAAG---TTGCATCACTAGTGGTCTTCTTCTTGTGTGTCTACCAG  715

Qy   761  CTGCTTCATACACTACTGGTGCCAGCTTACTGGTTGATGGTGGTTTTCACT  810
          | ||||||    |  |||||||||||||| ||||||||||||||||||||||
Db   716  CAGCTTCTTTATATTACTGGTCAAAACCATTTGTATTGATGGTGGTTTTTCACT  765
```

Search completed: March 2, 2005, 14:53:06
Job time : 564 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: March 2, 2005, 05:24:38 ; Search time 3855 Seconds
                                              (without alignments)
                                              10256.668 Million cell updates/sec

Title: US-10-720-018-1
Perfect score: 816
Sequence: 1 atgactgactacattccaac..........atggtggtttcacttcttgg 816
Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 2422607955 residues

Total number of Hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :   GenEmbl:*
                1: gb_ba:*
                2: gb_htg:*
                3: gb_in:*
                4: gb_om:*
                5: gb_ov:*
                6: gb_pat:*
                7: gb_ph:*
                8: gb_pl:*
                9: gb_pr:*
               10: gb_ro:*
               11: gb_sts:*
               12: gb_sy:*
               13: gb_un:*
               14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
	816	100.0	816	8	AJ583159	AJ583159 Ambrosioz
1	332.8	40.8	1166	8	CR382133_10	Continuation (11 o
2	305.6	37.5	1614	8	CTARDH	U00675 Candida tro
3	275.2	33.7	110000	8	YSAARDH	L16227 Candida alb
c 4	268.4	32.9	1157	8	CR382133_10	Continuation (11 o
5	238.2	29.2	696	6	PSARDH	Z46866 P.stiptis
6	238.2	29.2	696	6	AR031556	AR031556 Sequence
7	238.6	26.8	846	6	AR345028	AR345028 Sequence
8	218.6	19.1	110000	8	AR544898	AR544898 Sequence
9	155.8	16.0	110000	8	CR382121_09	Continuation (10 o
c 10	130.8	11.2	110000	8	CR382132_03	Continuation (4 of
c 11	92.6	11.3	110000	8	CR382133_11	Continuation (12 o
c 12	91.4	11.2	110000	8	CR382137_06	Continuation (7 of
13	87.2	10.7	4293	8	AF002134	AF002134 Candida a
c 14	86.8	10.6	126105	6	AC105434	AC105434 Magnaport
15	86.6	10.6	927	6	AR547359	AR547359 Sequence
16	81.8	10.0	1485	8	AY387591	AY387591 Emericell
c 17	81.8	10.0	15779	8	AY434691	AY434691 Emericell
c 18	78.6	9.6	110000	8	CR382125_21	Continuation (22 o
c 19	78.6	9.6	110000	8	CR382131_06	Continuation (7 of
c 20	72	8.8	110000	1	AE016827_14	Continuation (15 o
21	70.8	8.7	110000	8	CR388133_10	Continuation (11 o
22	70	8.6	1063	8	UFU81790	U81790 Uromyces fa
c 23	68.6	8.4	409	11	KLAJ9366	AJ293366 Kluyverom
c 24	68.2	8.3	340900	1	SMES91791	AJ591791 Sinorhizo
c 25	67.6	8.3	32814	8	SPAC22A12	Z99295 S.pombe chr
c 26	64	7.8	6430	8	AFU320167	AJ320167 Aspergill
c 27	64	7.8	10126	1	CR382138_00	Continuation (1 of
28	63.6	7.8	309135	1	AE001711	AE001711 Thermocog
29	62.4	7.6	873	6	AP005946	AP005946 Bradyrhiz
30	61.8	7.6	786	6	AR547523	AR547523 Sequence
c 31	61.4	7.5	110000	8	CR382132_02	Continuation (3 of
32	61	7.5	822	6	AJ245634	AJ245634 Solanum t
33	59.8	7.3	1129	8	AX772916	AX772916 Sequence
c 34	59.8	7.3	15350	8	DURTROPRI	L20473 Datura stra
35	59.4	7.2	976	8	SPACBE11	AL021817 S.pombe c
c 36	59	7.2	774	6	E30144	E30144 Arabinitol
37	59	7.2	2255	6	E30145	E30145 Arabinitol
38	57.8	7.1	300787	1	AE016744	AE016744 Staphyloc
39	57.2	7.0	37969	8	CR382137_20	Continuation (21 o
40	57.2	7.0	110000	8	CR382137_19	Continuation (20 o
41	56.6	6.9	976	8	STU400815	AJ400815 Solanum t
42	56.4	6.9	789	8	BT010425	BT010425 Arabidops
43	56.4	6.9	825	6	AX772918	AX772918 Sequence
44	56.4	6.9	909	6	AK175221	AK175221 Arabidops
45	56.4	6.9	1099	8	D88156	D88156 Hyoscyamus

ALIGNMENTS

RESULT 1
AJ583159

| LOCUS | AJ583159 | 816 bp | mRNA | linear | PLN 13-APR-2004 |

DEFINITION Ambrosiozyma monospora partial mRNA for NADH L-xylulose reductase
            (alxl gene).
ACCESSION   AJ583159
VERSION   AJ583159.1 GI:42820317
KEYWORDS   alxl gene; NADH L-xylulose reductase.
SOURCE   Ambrosiozyma monospora
  ORGANISM Ambrosiozyma monospora
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycopsidaceae; Ambrosiozyma.
REFERENCE   1
  AUTHORS   Verho,R., Putkonen,M., Londesborough,J., Penttila,M. and Richard,P.
  TITLE   A Novel NADH-linked L-Xylulose Reductase in the L-Arabinose
            Catabolic Pathway of Yeast
  JOURNAL   J. Biol. Chem. 279 (15), 14746-14751 (2004)
   PUBMED   14736891
REFERENCE   2 (bases 1 to 816)
  AUTHORS   Richard,P.
  TITLE   Direct Submission
  JOURNAL   Submitted (18-SEP-2003) Richard P., VTT Biotechnology, VTT, PO box
            1500, Tietotie 2, Espoo, 02044 VTT, FINLAND
FEATURES   Location/Qualifiers
     source   1..816
                     /organism="Ambrosiozyma monospora"
                     /mol_type="mRNA"
                     /strain="NRRL Y-1484"
                     /db_xref="taxon:43982"
     gene   1..816
                     /gene="alxl"
     CDS   1..>816
                     /gene="alxl"
                     /codon_start=1
                     /product="NADH L-xylulose reductase"
                     /protein_id="CAE47547.1"
                     /db_xref="GI:42820318"
                     /translation="MTDYIPTFRFDGHLTIVTGACGGLAEALIKGLLAYGSDIALLDI
                     DQEKTAAKQAEYHKYATEELKLKEVPKMGSYACDISDSDTVHKVFAQVAVKDGKLPLH
                     LVNTAGYCENFPCEDYPAKNAEKMVKVNLLGSLYVSQAFAKPLIKEGIKGASVVLIGS
                     MSGAIVNDPQNQVYYNSKAGVIHLAKTLACEWAKYNIRVNSLNPGYIYGPLTKNVIN
                     GNEELYNRWISGIPQQRMSEPKEYIGAVLYLLSESAASYTTGASLIVDGGFTSW"

RESULT 2

LOCUS   CTARDH   1166 bp   DNA   linear   PLN 16-MAY-1995
DEFINITION Candida tropicalis D-arabinitol dehydrogenase (ARD) gene, complete
            cds.
ACCESSION   U00675
VERSION   U00675.1 GI:392785
KEYWORDS   .
SOURCE   Candida tropicalis
  ORGANISM Candida tropicalis

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE   1 (bases 1 to 1166)
  AUTHORS   Murray,J.S., Wong,M.L., Miyada,C.G., Switchenko,A.C., Goodman,T.C.
            and Wong,B.
  TITLE   Isolation, characterization and expression of the gene that encodes
            D-arabinitol dehydrogenase in Candida tropicalis
  JOURNAL   Gene 155 (1), 123-128 (1995)
   PUBMED   7698655
  MEDLINE   95212917
REFERENCE   2 (bases 1 to 1166)
  AUTHORS   Wong,B.
  TITLE   Direct Submission
  JOURNAL   Submitted (10-AUG-1993) Brian Wong, Internal Medicine, University
            of Cincinnati College of Medicine, 231 Bethesda Avenue, Cincinnati,
            OH 45267-0560, USA

FEATURES   Location/Qualifiers
     source   1..1166
                     /organism="Candida tropicalis"
                     /mol_type="genomic DNA"
                     /strain="ATCC 750"
                     /db_xref="taxon:5482"
                     /clone="pCT1A"
                     /clone_lib="genomic DNA library in pAB107 from V.F. Kalb
                     and J.C. Loper"
     gene   241..1089
                     /gene="ARD"
     CDS   241..1089
                     /gene="ARD"
                     /function="catalyzes D-arabinitol + NAD <=> D-ribulose +
                     NADH"
                     /codon_start=1
                     /evidence=experimental
                     /product="D-arabinitol dehydrogenase"
                     /protein_id="AAA66355.1"
                     /db_xref="GI:392786"
                     /translation="MDSSSYNSYDNIVPSFRLDGKLVIITGGSGGLSAVVSRALLAKG
                     ADIALIDMNLERTQQAARDVLQWGEEQMKGKHESPIGQVSAWSCNIGDAEAVELTFKA
                     INEHGKVASVIINTAGYAENFPAEEYPAKNAENIMKVNGLGSFYVSQAFARPLIQNN
                     MTGSIILIGSMSGTIVNDPQPQCMYNMSKAGVIHLARSLACEWAKNIRVATLSPGYI
                     LTPLTRNVISGHTEMKTEWESKIPMKRMAEPKEFVGSSILYLASDSASSYTGHNLVVD
                     GGYEWCW"
ORIGIN

Qy   Query Match   40.8%; Score 332.8; DB 8; Length 1166;
           Best Local Similarity   65.0%; Pred. No. 1.7e-73;
Matches   528; Conservative   0; Mismatches 272; Indels 12; Gaps 2;

Qy   14 TTCCAACTTTAGATTCGATGGCCACTTAACCATTGTCACAGGTGCCATGC 73
              |||||| ||||||||||| |||| ||||| |||||| |||||| ||||
Db   278 TTCCAAGTTTTAGATTGGATGGAAAATTAGTAATATATCACTGGTTGGTTTGT 337

Qy   74 CTGAAGCTTTAATCAAGGTGTTTTGTTGGCCTACCGGTTCTTGACATTGCTTTGTGATATCG 133
              |||||||||||| |||||| ||| |||||||||||||||| ||| |||||| |||| |||||
Db   338 CCCGCTGTTGTTCCAAACAAGGTGTCGTTCTCGACATTGCTTTGTGATATCG 397

Qy   134 ACCAAGAAAAAGACTGCTGCCAAACAAGCCGAATACCACAAATACCGCTACTGAAGAATTGA 193
              ||| ||||||| |||||||||||||||||||||||||||| ||||||| || ||||| |
Db   398 ACTTGGAAAGAGACCCAACAAGCCGCAGAGATGTTTACAATGGGTGAAGAACAAATGA 457

Qy   194 AGTTGAAAGA----------------AGTTCCAAAGATGGGTGTTCATATGCCTGTGATATTTCTGATT 244
              ||| |||||| ||||||| |||||||||||| ||||| |||||||| |||||
Db   458 AAGGAAAAACACGAATCTCCAATTGGTCAAGTCAGTGCTTGGTCCTGTAACATTGGTGATG 517

Qy   245 CTGATACCGTTCACAAGGTAGTTTTGCTCAAGGTTGCTAGGAGTTTTGGTAAGTTGCCATGC 304
              ||| |||||||||||||| ||| ||||||||||||||||| |||||||||||||||| |||
Db   518 CTGAAGCTGTTGAATTGACATTCAAAGCCATCAACGAACACGGTAAAGTTGCCAGTG 577

Qy   305 ACTTGGTTAACACAGCTGGTTACTGTGAAAACTTCCCATGTGAAGAATTACCCAGGTTG 364
              |||||||||| | |||||||| |||||||||||| ||||||| |||||||||||||
Db   578 TTTTGATTAACACAGCTGGTTATGCTGAAAATTTCCCTGCTGAAGAATACCCAGCCAAGA 637

Qy   365 ACGGTGAGAAGATGGTGAAGGTTAACTGTTGGGGTTCTTGTGTGGTTTCTCAAGCCTTTG 424

RESULT 3

```
Db   638  ATGCCGAAAATATCATGAAGGTCAAATGGTTTAGGATCCTTTAGTCTCACAAGCTTTCG   697
          |||| |||                         |||| |  |||||   ||||| |||
Qy   425  CTAAGCCATTGATCAAAGAAGGTATCAAGGGTGCTTCTCGTTGTTTGATTGGTTCTATGT  484

Db   698  CCAGACCTTTGATTCATCAAAA---TAACACTACGGCTCTATTATACTTGATTCGTTATGT  754
          || | |||||||| |  ||||||    |||||||| ||||| ||| |||||| || ||||
Qy   485  CTGGTGCCATTGTCAAGGATCCTCAAAAACCAAGTGTTCTACAACATGTCCAAGGCTGGTG  544

Db   755  CTGGTACCATTGTCAACGATCCACAACCAAGTATGTACACACATGTCCAAAGCTGGTG     814
          |||||  ||||||| || |||||| ||   |||  |||| ||||||||||  ||||||
Qy   545  TTATCCATTTGGCTAAGACTTTGGCTTGTTGAATGGGCTAAGTACAAATATCAGAGTTAATT 604

Db   815  TCATTCACTTGGCCAGATCTTTAGCATGTGAATGGGCTAAATACAACATCAGAGTCAATA   874
          | || |||||||  || |||| | |  ||||||||||| ||| ||||| ||||||| |||
Qy   605  CTTTAAACCCAGGTTACATCTACGGTCCTCCTTTGACCAGAAATGTTATCAATGGTCACAT  664

Db   875  CCTTGTCCCAGGTTATATTTTGACTCCATTGACAAGAAATGTTATTAGTGGTCACAT     934
          ||||| ||| |||||||| || |||||| |||  ||||||||||||  |||||||||
Qy   665  AATTGTACAACAGATGGATCTCTGGTATCCCACAACAAACAAAGAAATGTCCGAACCAAAGGAAT 724

Db   935  AAATGAAGACAGAATGGGAAATCCAAGATTCCAATGAAGAGAATGGCCGAGCCAAAAGAAT   994
          |||||  |||||||||||| | ||| ||| |||| |||||| ||||||| |||| |||||
Qy   725  ACATTGGTGCTGTTTTGTTACTTGCTTTCATACACTACTACTGGTGCCA              784

Db   995  TTGTCGGTTCCTATTTGTACTGCCTTCTTCTTCCTACACTACTGGTCACA            1054
          |||| ||||| | |||||||| || ||| |||| | |||||||||| |||
Qy   785  GCTTACTCGGTTGATGGTCGGTTTCACTTCTTGG                            816

Db  1055  ATTTGGTTAGTGGATGGAGGTGGTTACCGAATGTTGG                        1086
```

LOCUS   YSAARDH   1614 bp   DNA   linear   PLN 21-JUN-1995
DEFINITION   Candida albicans D-arabinitol dehydrogenase (ArDH) gene, complete cds.
ACCESSION   L16227
VERSION   L16227.1   GI:295567
KEYWORDS   ArDH gene; D-arabinitol dehydrogenase.
SOURCE   Candida albicans
  ORGANISM   Candida albicans
      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
      Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1 (bases 1 to 1614)
  AUTHORS   Wong,B.; Murray,J.S.; Castellanos,M. and Croen,K.D.
  TITLE   D-arabinitol metabolism in Candida albicans: studies of the
           biosynthetic pathway and the gene that encodes NAD-dependent
           D-arabinitol dehydrogenase
  JOURNAL   J. Bacteriol. 175 (19), 6314-6320 (1993)
  MEDLINE   94012494
  PUBMED   8407803
REFERENCE   2 (sites)
  AUTHORS   Wong,B.; Leeson,S.; Grindle,S.; Magee,B.; Brooks,E. and Magee,P.T.
  TITLE   D-arabinitol metabolism in Candida albicans: construction and
           analysis of mutants lacking D-arabinitol dehydrogenase
  JOURNAL   J. Bacteriol. 177 (11), 2971-2976 (1995)
  MEDLINE   95286472
  PUBMED   7768790
REFERENCE   3 (bases 1 to 1614)
  AUTHORS   Wong,B.
  TITLE   Direct Submission
  JOURNAL   Submitted (18-MAY-1993) Brian Wong, Department of Internal
           Medicine, University of Cincinnati College of Medicine, OH
           45267-0560, USA
COMMENT   Original source text: Candida albicans (strain WO1) DNA.
FEATURES   Location/Qualifiers
  source   1..1614
           /organism="Candida albicans"
           /mol_type="genomic DNA"
           /strain="WO1"
           /db_xref="taxon:5476"
           /map="6; segment 0"
  gene   435..1280
           /gene="ArDH"
  CDS   435..1280
           /gene="ArDH"
           /function="catalyzes D-arabinitol + NAD = D-ribulose + NADH"
           /codon_start=1
           /evidence=experimental
           /transl_table=12
           /product="D-arabinitol dehydrogenase"
           /protein_id="AAC37430.1"
           /db_xref="GI:295568"
           /translation="MDSAYWSYDNIVPSFRLDGKLVIITGGSGGLAAVVSRALLAKGA
           DVALVDMNLERTCQAARDVLQWGEEQMKGKYESPIGQVSAWSCNIGDAEAVDLTFKAI
           NEHHGKISSVLVNTAGVAENFPAEEYPAKNAENLMKVNGLGSFYVSQAFARPLIQNNM
           TGSIILLGSMSGTIVNDPQPOCMYNMSKAGVIHLARSLACEWAKYNIRVNTLSPGYIL
           TPLTRNVISGHTEMKTEWESKIPMKRMAEPKEFVGSIILYLASESSASSYTTGHNLVVDG
           GYECW"

ORIGIN

Query Match 37.5%; Score 305.6; DB 8; Length 1614;
Best Local Similarity 62.9%; Pred. No. 1.3e-66;
Matches 511; Conservative 0; Mismatches 289; Indels 12; Gaps 2;

```
Qy    14  TTCCAACTTTTAGATTCGATGGCCACTTAACCATTGTCACAGGTCCGCTGGTTAG       73
          |||||   |||||   ||| ||||      ||||||| ||| ||||| ||||||
Db   469  TCCCAAGCTTCCGTTTGGATGGAAAAC---------AGATGGGGTTCATAATGCCTTGCGATG  528

Qy    74  CTGAAGCTTTAATCAAAGGTTTGTTGGCCTTACGGTTCTGACATTGCTTTGCTTTGATCG  133
          |  ||||||||| ||||||||||||||| |||   ||| ||||| ||  |  | ||| ||
Db   529  CTGCCGGTGGTATCAAGAGCTTTATTAGCCCAAAAGTTGCCGGATGTTGCATTAGTCGGATATGA  588

Qy   134  ACCAAGAAAAGACTTGCCAAACGAAGCCGGAATACCACAAACATACGCTACTGAAGAATTGA  193
          |     |||||   |||| ||  |||| ||  |  |||||| | | | | | |||| | ||
Db   589  ACTTGGAAAGAACAACAAGCTGCTAGAGCCGTCTTACAATGCGGGCGAGAGCAAATGA      648

Qy   194  AGTTGAAAGAAGTTCCAA-----------AGATGGGGTTCATATGCCTGATATTTCTGATT  244
          |||  |||  ||||||||           |  ||| | | ||||| ||| | ||||| |
Db   649  AAGGTAAATACGAATCACCAAATCGGTCACGTGAGTGCTTCGGTCATGTTAATATTGCGATG  708

Qy   245  CTGATACCGTTCACAAGGTGTTTGCTCAAGTTGCTAAGGATTTTGGTAAGTTGCCATTGC  304
          |||| | | |||| ||| |||| ||||||   ||| | | ||| ||| ||||||| | |
Db   709  CTGAAGCTGTGTCGACTTGACATTCAAAGCCATCAACGAACCACGGCCAAAATCTCAAGTG  768

Qy   305  ACTTGGTTAACACAGCCTGGTTACTGTTGAAAACTTCCCATGTGAAGATTACCAGCCAAGA  364
          |||||||| ||||    ||| | ||||| ||||||| | |||| | || ||||||| ||
Db   769  TCTTGGTCAACACTGCCGGTGTTGCCGAAAACTTCCCAGCTGAAGAGTACCCAGCCAAGA  828

Qy   365  ACGGTGAGAGAATGGTGAAGGTTAACTGTTGGTGATGTTTCCAAGCTTTTG          424
          |  || ||||| ||| || |||||| |||| ||||| || ||||| ||||
Db   829  ACGCTGAAAACCTTATGAAAGTTAACGGTTTGGGTTCATTCTACGTTTCCCAAGCTTTTG  888

Qy   425  CTAAGCCATTGATCAAAGAAGGTGCTTCTCGTTGTTTGATTGGTTCTATGT          484
          | |  ||||||| |||| || | |  ||| ||| |  | ||| | ||||
Db   889  CTAGACCATTAATCCAAAA----CAACATGACCGGATCAATCATTTTGATCGGGTCAATGT  945

Qy   485  CTGGTGCCATTGTCAAGGATCCTCAAAAACCAAGTGTTCTACAACATGTCCAAGGCTGGTG  544
          ||||||||| | ||   ||||| |||| |||||||| | |||| ||||| ||||| |||||
Db   946  CCGGTACCATCGTCAACGACCCACAACCAAGTGTACAACACATGTCAAACGCGGTG        1005

Qy   545  TTATCCATTTGGCTAAGACTTTGGCTTGTTGAATGGGCTAAGTACAAATATCAGAGTTAATT  604
          ||||||||| | || | | |||||||||| |||||| || | ||||| ||||||| |||
Db  1006  TCATTCATTTAGCCAGATCTTTGTGTGAATGGGCTAAATACAACATCAGAGTCAACA        1065

Qy   605  CTTTAAACCCAGGTTACATCTACGGTCCTCCTTTGACCAGAAATGTTATCAATGGTCACAT  664
          |||| |  ||| ||||| || |||| |  || |||  ||| ||||| ||  ||||||||
Db  1066  CATTGTCGCGCCAGGGTTACATCTTAACCCCATTGACAAGAAATGTTATTAGTGGACACCG  1125

Qy   665  AATTGTACAACAGATGGATCTCTGGTATCCCACAACAAACAAAGAAATGTCCGAACCAAAGGAAT 724
          |||||  | ||||| || |||  ||| | | || ||| ||| |||||| |||  ||||||||
Db  1126  AAATCAAGACAGAATGGGAATCAAAGACTCCAATGAAGAGAATGGCAGAGAACCAAAAGAGT  1185
```

```
Qy    725  ACATTGGTGCTGTGTTTGTACTTGCTTCGTTCTGAAATCGTCTCATACACTACTGGTGCCA  784
           |||| ||  || | |  | ||||| |  |  | ||| ||| |||| |||||||||||| ||
Db   1186  TTGTGGTTCTATCTATATACTTGGCCCTCAGAAATCGCTTCATCATACACTACTGGTCACA  1245

Qy    785  GCTTACTCGGTTGATGGTGGTTTCACTTCTTGG  816
           ||| | | |||||| || ||| || ||| ||||
Db   1246  ACTTGGTCGTTGACGGGGGGTTACGAGTCTTGG  1277
```

RESULT 4
CR382133_10/c
WPCOMMENT

Sequence split into 13 fragments

Fragment Name	Begin	End
CR382133_00	1	110000
CR382133_01	100001	210000
CR382133_02	200001	310000
CR382133_03	300001	410000
CR382133_04	400001	510000
CR382133_05	500001	610000
CR382133_06	600001	710000
CR382133_07	700001	810000
CR382133_08	800001	910000
CR382133_09	900001	1010000
CR382133_10	1000001	1110000
CR382133_11	1100001	1210000
CR382133_12	1200001	1249565

LOCUS CR382133 Accession CR382133

```
Qy     14  TTCCAACTTTTAGATTCGAATGGCCATCAACCATTGTCACAGGTGCCTGTGGTGGTTTAG   73
Db  28160  TTCCAAGCTTTAGATTAGGATGGTAGATTAGCTAGTACTTCTTTGCTTCTTAG        28101

Qy     74  CTGAAGCTTTAATCAAGGGTTGTTGGCCTACGGTTCTCGACATTGTCTTTGCTTTGATATCG  133
Db  28100  CTGCTGTGTCTAGACCATTATTAGTCTCAGGGAGCACAAATTGCTTTGCTTTGATATGA    28041

Qy    134  ACCAAGAAAAGACTGCTTGCCAAACAACAGCCGAATACCACACAAAAATACCTACTGAAGAATTGA  193
Db  28040  ATTTAGAAAAGAACCAAGGCTGCAGCAAAGGGAATTGGAAACATGGGGACAAAGAAATCATTGA  27981

Qy    194  AGTTGAAAGA------------AGTTCCAAAGATGGGTTCATAGTGCCTGTGTATTTCTGATT  244
Db  27980  AAGGTGCACATGAAAGCCCAGTGGGGAGCTGTTTCCGCATGGGTCTTGTGTAATATTGGGGGATT  27921

Qy    245  CTGATACCGTTCACAAGGTGTTTGCTCAAGTGTCTAAGGATTTTGGTAAGTTGCCCATTGC  304
Db  27920  TTGAACAGGTGGAAGAGTGCTTTAAGAACATCAATGAACACCACGATATGGTGCTCGATT  27861

Qy    305  ACTTGGTTAACACAGCTGGTTACTGGTGAAAACTTCCCATGTGAAGAAGAATCATTGA    364
Db  27860  TGTTAATCAATACTGCAGGTTACTGTGAGAATTTCCCGCGAAGAATACCCATCTGCCA  27801

Qy    365  AGGCTGAGAGATGGTGAAGGTTAACTTGTTGTGGGTTCTTGTATGTTTCTCAAGCCCTTTG  424
Db  27800  ATGCTGAAGGGTATTTGAAGGTCAAATGGAATGGGTCTTGTGCATTCACAAGCATTTG   27741

Qy    425  CTAAGCCATTGATCAAAGAAGGTATCAAGGGTGCTTCTGTTGTTTTGATTGGTTCTATGT  484
Db  27740  CCCGTCCATTGATCA---GCAGCAAATAAGAAAAGGATCCATCATTTAGTTGGGTCGATGT  27684

Qy    485  CTGGTGCCATTGTCAACGGATCCTCAAAAACCAAGTGTCTACAACAGTTCCAAGGCTCGGTG  544
Db  27683  CTGGAACAATTGTGAACCCACCACCACAATGTATGTACAACATGTCCAAGGCAGGTG   27624

Qy    545  TTATCCATTTGGGTAAGACTTTGGCTTGTGAATGGGCTAAGTACAACATCAGAGTTAATT  604
Db  27623  TTATTCACTTGACCAGAATCATTGGCTTGTTTGAAATCAACATCAAGAGTTAACA     27564
```

Query Match 33.7%; Score 275.2; DB 8; Length 110000;
Best Local Similarity 60.6%; Pred. No. 6e-59;
Matches 492; Conservative 0; Mismatches 308; Indels 12; Gaps 2;

Continuation (11 of 13) of CR382133 from base 1000001 (CR382133 Debaryomyces hansenii ch...

```
Qy    605  CTTTAAACCCAGGTTACAYCTACGGTCCTCTTTTGAACCAAGAAATGTTATCAATGGTAAGGAG  664
Db  27563  CTTTGAGTCCTGGATATATCTTAACCCCTTTGACCAAGAAACGTCATTTCTTGGTGTTCCG     27504

Qy    665  AATTGTACAACAGATGGATCTCTCGGTATCTCCACAACACAGAAGAAATGTCCGAAAGGAAT  724
Db  27503  ATATGAAGGAAAGCATGGGAATCAAAAGGTCCCAATGAAGAGAAGAAATGGCCCAAACCAAAAGGGAAT  27444

Qy    725  ACATTGGTGCTGTGTTTGTACTTGCTTCGTTCTGAAATCGTCTCATACACTACTGGTGCCA  784
Db  27443  TTGTGGCTCGATCTTGTACTTGGCCACCGGAGTCTGCTTCTTAGTTATACTACAGGTCACA  27384

Qy    785  GCTTACTCGGTTGATGGTGGTTTCACTTCTTGG  816
Db  27383  ACTTAGTTGTTGATGGTGGCTACCGAGTGCTGG  27352
```

RESULT 5
PSARDH
PSARDH

LOCUS   PSARDH   1157 bp   mRNA   linear   PLN 29-MAR-1996
DEFINITION P.stipitis ARDH gene encoding D-arabitol dehydrogenase.
ACCESSION   Z46866
VERSION   Z46866.1 GI:758241
KEYWORDS   D-arabitol dehydrogenase.
SOURCE   Pichia stipitis
  ORGANISM Pichia stipitis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE   1 (bases 1 to 1157)
  AUTHORS   Hallborn,J., Walfridsson,M., Penttila,M., Keranen,S. and
            Hahn-Hagerdal,B.
  TITLE   A short-chain dehydrogenase gene from Pichia stipitis having
            D-arabinol dehydrogenase activity
  JOURNAL   Yeast 11 (9), 839-847 (1995)
  MEDLINE   96090133
   PUBMED   7483848
REFERENCE   2 (bases 1 to 1157)
  AUTHORS   Hallborn,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (30-NOV-1994) Johan Hallborn, Applied Microbiology, Lund
            Institute of Technology, Chemical center, Getingevagen 60, Lund,
            S-221 00, Sweden
FEATURES   Location/Qualifiers
     source   1..1157
                     /organism="Pichia stipitis"
                     /mol_type="mRNA"
                     /strain="CBS 6054"
                     /db_xref="taxon:4924"
     gene   1..1157
                     /gene="ARDH"
     CDS   191..1027
                     /gene="ARDH"
                     /function="NAD dependent, oxidises D-arabitol to
                     D-ribulose"
                     /codon_start=1
                     /evidence=experimental
                     /product="D-arabitol dehydrogenase"
                     /protein_id="CAA86939.1"
                     /db_xref="GI:763164"
                     /db_xref="GOA:P50167"
                     /db_xref="UniProt/Swiss-Prot:P50167"
                     /translation="MDYSYANVVPNRLDGRLAIITGGSGGLAAVISRALLAQGADVA
                     LIDMNLERTKSAAKEVLGWGEETLKGEHASAIGQVSAWSCNIGDAEAVDATFSSINEH
                     HGKIADLLINFAGYCENPPAETYPATNAESIMKVNLGSFVSQSFARPLIQONNLRGS
                     ILLIGSMSGTIVNDPQPCMYNMSKAGVIHLVRSLACEWAKYNIRVNTLSPGYILTPL
                     TRNVISGHTEMKEAWESKIPMKRMAEPKEFVGSILIVLASETPASSYTTGHNLVVDGGYE
                     CW"
ORIGIN
```

Query Match 32.9%; Score 268.4; DB 8; Length 1157;
Best Local Similarity 60.1%; Pred. No. 3.6e-57;
Matches 487; Conservative 0; Mismatches 311; Indels 12; Gaps 2;

## RESULT 6

```
LOCUS        AR031556      696 bp    DNA    linear    PAT 29-SEP-1999
DEFINITION   Sequence 6 from patent US 5866382.
ACCESSION    AR031556
VERSION      AR031556.1   GI:5945845
KEYWORDS     .
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 696)
  AUTHORS    Hallborn,J., Penttila,M., Ojamo,H., Walfridsson,M., Airaksinen,U.,
             Keranen,S. and Hahn-Hagerdal,B.
  TITLE      Xylose utilization by recombinant yeasts
  JOURNAL    Patent: US 5866382-A 6 02-FEB-1999;
FEATURES             Location/Qualifiers
  source    1..696
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN

Query Match   29.2%;  Score 238.2;  DB 6;  Length 696;
Best Local Similarity   64.4%;  Pred. No. 1.7e-49;
Matches  373;  Conservative   0;  Mismatches  203;  Indels   3;  Gaps   1;
```

```
Qy     5 CTGACTACATTCCAACTTTTAGATTCGATGGGCCACTTAACCATTGTCTCACAGGTGCCTGTG 64
Db   207 CTAACGTTCGTTCCCAACTTCAGATTGGACGGAAGATTGGCTATTATTACCGGAGGTTCTG 266

Qy    65 GTGGTTTAGCTGAAGCTTTAAATCAAGGGGTTTGTTGGGCCTTACGGTTCTGACATTGCTTTGC 124
Db   267 GTGGTTTGGCCGCAGTCATTTCCCGTGCCTTGTTGTCGGCGCGGCTGAGTGTGCTCTCA 326

Qy   125 TTGATATCGACCAAGAAAAAGACTGCTGCCAAACAAGCCGAATACCACATAACGGCTACTG 184
Db   327 TTGACATGAACTTGGAAACAAGACCAAGTCCCGTCCCTGCCAAAGAAGTTTGGGCTGGAAG 386

Qy   185 AAGAAATTGAAAGTTGAAAGAAGTTCCAAAGATGGGTTCATATG 235
Db   387 AGACGTTGAAAGGGTGAACACGGCTTCAGCCATCGGCCAAGTTTCCGCCCTGGTCCTCCAACA 446

Qy   236 TTTCTGATTCTGATACCGTTCACAAGGTGTTGCTAAGGGTTCATATG------------CCTGTGATTA 295
Db   447 TTGGGGATGTCGAGCGAGTAGACGCTACTTTCAGCTCCATCAACGAACACCGGCCAAGA 506

Qy   296 TGGCCATTGCACTTGGTTAACACAGCTGGTTACTGTGAAGAATACC 355
Db   507 TCGCTGACTTGTTGATTAACACGCCTGGATACTCGTGAAAAAACGTACC 566

Qy   356 CACGCCAAGAAACGCTGAGAAGAGATGGTGAAGGTTAACTTGTTGGTTCTTTTGTATGTTTCTC 415
Db   567 CGGCTACTAACGCTGAAAGCATCATGAAGGGCTCATCTTCACGTTTTCGC 626

Qy   416 AAGCCTTTGCTAAGCCATTGATCAAAGAAGGTATCAAGGGGTATCAAGTTACC 475
Db   627 AATCGTTCGCTACGACCATTGATCCAGAA---CAACTTGAGAGAGTTCATGATG 683

Qy   476 GTTCTATGTCTCGGTGGTCCATTGTCAAACAGTGTCTCTACAAACAGTTCCA 535
Db   684 GCTCAAATGTCTCGGAACAATGTCACACGACCCAATGTTAGTACAAACATGTCCA 743

Qy   536 AGGCTGGCGTGTTATCCCATTTGGGGGTTAAGACATTTGGGCTTAGGTACTACAACATCA 595
Db   744 AGGCTGGGACTTGGATCCACTTGGTGGGCTTGGCCTGCAGTCGGTTACGAACATCA 803

Qy   596 GAGTTAATTCTTTAAACCCAGGTTCTTAGACGGTCCTTGACCAAGAATGTTATCAATG 655
Db   804 GAGTCAACACCTTATCACCAGGCTTATATTTGACTCTCTTAACCAGGAAACGTGATTTCTG 863

Qy   656 GTAACGAACGAAATTGTACAACAGAGATCTCTGGTATCCCACAACAAAGAGTTCCGAC 715
Db   864 GCCACACAGAGATGAAGAGCCTGGGAATCCCATGAAGAGAGAATGGCCGAAC 923

Qy   716 CAAAGGAATACATTGGGTGCTGTTGTTTGTACTTGCTTCTTCTGCTTCATTACACTA 775
Db   924 CCAAGGAATTCGTGGGGTCCATCTTTATACTTGGCAAGCGACTGCTTCTTCCTACACTA 983

Qy   776 CTGGTGCCAGCTTACTCGGTTGATGGTGGTT 805
Db   984 CGGGCCACAAATTTGGTTGTGGACGGAGGAT 1013
```

RESULT 7

```
LOCUS        AR345028      696 bp    mRNA    linear    PAT 17-AUG-2003
DEFINITION   Sequence 6 from patent US 6582944.
ACCESSION    AR345028
VERSION      AR345028.1   GI:33741148
KEYWORDS     .
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 696)
  AUTHORS    Hallborn,J., Penttila,M., Ojamo,H., Walfridsson,M., Airaksinen,U.,
             Keranen,S. and Hahn-Hagerdal,B.
  TITLE      Production of ethanol from xylose
  JOURNAL    Patent: US 6582944-A 6 24-JUN-2003;
FEATURES             Location/Qualifiers
  source    1..696
            /organism="unknown"
            /mol_type="mRNA"
ORIGIN

Query Match   29.2%;  Score 238.2;  DB 6;  Length 696;
Best Local Similarity   64.4%;  Pred. No. 1.7e-49;
```

```
Qy   227 CCTTGTGATATTTCTGATTCTGATACCGTTCACAAGGTGTTGCTAAGGGTT 286
Db   107 CCTGCAACATTGGGGATGTCGAGCGAGTAGACGCTACTTTCAGCTCCATGTG 166

Qy   287 TTGGTTAAGTTGCCATTGCACTTGGTTAACACAGCTGGTTACTGTGAAAAACTTCCCATGTG 346
Db   167 ACGGCCAAGATCGCTGACTTGTTGATTAACACGCCTGGATACTCGTGAAAAACG 226

Qy   347 AAGATTACCCAAGGACGACGCTGAGAAGAGATGGTGAAGGTTAACTTGTTGTCTACA 406
Db   227 AAACGTTACCCCGGGTACTTACCTAACGGCATGATCATGAAAGAAGGTGAACCTCAATTCT 286

Qy   407 ATGTTTCTCAAGCCTTTGCTAAGGCCATTGATCAAAGAAGGTATCAAGGGG 466
Db   287 ACGTTTCGCAATCGTTCGCTACGACCATTGATCCAGAA---CAACTTGAGAGAGTTCATCA 343

Qy   467 TTTTGATTGGTTCTATGTCTCGGTGGCCATTGTCAAACAGTGTCTCTACA 526
Db   344 TCTTGATTGGCTCAAATGTCTCGGAACAATGTCAACGACACCAAGTTGTCTACA 403

Qy   527 ACATGTCCAAGGCTGGTGTGTTATCCCATTTGGCTAAGACTTGGG 586
Db   404 ACATGTCCAAGGCTGGGACTTGGATCCACTTGGTGGGCTTGGCCTGCAGT 463

Qy   587 ACAACATCAGAGTTAATTCTTTAAACCCAGGTTCTTACGGTCCTTGACCAAGAATG 646
Db   464 ACAACATCAGAGTCAACACCTTATCACCAGGCTTATATTTGACTCTCTTAACCAGGAAACG 523

Qy   647 TTATCAATGGTAACGAACGAAATTGTACAACAGAGATCTCTGGTATCCCACAACAAAGAA 706
Db   524 TGATTTCTGGCCACACAGAGATGAAGAGCCTGGGAATCCCATGAAGAGAGAA 583

Qy   707 TGTCCGAACCAAAAGGAATACATTGGGTGCTGTTGTTTGTACTTGCTTCTTCTGCTTCATTACACTA 766
Db   584 TGGCCGAACCCAAGGAATTCGTGGGGTCCATCTTTATACTTGGCAAGCGACTGCTTCTT 643

Qy   767 CATACACTACTGGTGCCAGCTTACTCGGTTGATGGTGGTT 805
Db   644 CCTACACTACGGGCCACAAATTTGGTTGTGGACGGAGGAT 682
```

Query Match   29.2%; Score 238.2; DB 6; Length 696;
Best Local Similarity   64.4%; Pred. No. 1.7e-49;

```
Matches 373;    Conservative    0;    Mismatches    203;    Indels    3;    Gaps    1;

RESULT 8
LOCUS       AR544898
AR544898
DEFINITION  Sequence 29 from patent US 6747137.
ACCESSION   AR544898
VERSION     AR544898.1  GI:5393073
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Weinstock,K.G. and Bush,D.
  TITLE     Nucleic acid sequences relating to Candida albicans for diagnostics
            and therapeutics
  JOURNAL   Patent: US 6747137-A 29 08-JUN-2004;
FEATURES             Location/Qualifiers
     source          1..846
                     /organism="unknown"
                     /mol_type="genomic DNA"

ORIGIN

Query Match    26.8%;  Score 218.6;  DB 6;  Length 846;
Best Local Similarity  62.0%;  Pred. No. 1.5e-44;
Matches 385;  Conservative 0;  Mismatches 224;  Indels 12;  Gaps 2;
```

```
Qy      227  CCTGTGATTATTTCTGATTCTGAATACCGTTCACAAGGTGTTGCTCAAGTTGCTAAGGATT  286
                ||| |||||||  || |||| | |||  ||   |||||||| || ||  |  ||   ||
Db      107  CCTGCAACATTGGGGATCGTGAGGCAGTAGACGCTACTTTCAGCTCCATCAACGAACACC  166

Qy      287  TTGGTAAGTTGCCATTGCACTGGTTAACACCGTTGACTTTGTGAAAACTTCCCATGTG    346
             |||| || ||  |||  |||||||| | || |||  ||||| ||| |||| |||  |
Db      167  ACGGCAAGATCGCTGACTTGTTGATTAACACCGGTCAGGATGCTTGGGTCAATGGGGTCCG  226

Qy      347  AAGATTACCCAGCCAAGAAACGCTGAGAAGAGATGGTGAAGGTTAACTTGTTGGTTCGTGT  406
             || || || |   ||| |  ||||| ||||| | ||||| ||| |||| ||    || ||
Db      227  AAACGTACCCGGTTACTAACGCTGAAAAGCATCATGAAAGCGGTTTTGGGGGCTCATTCT  286

Qy      407  ATGTTTCTCAAGCCTTTGCTCCATTGACCATTGATCAAAGAAGGTATCAAGGGTGCTTCGTTG  466
             |  |||  |  ||  |||| || ||  |   |  | | | |||   | |  |  ||| ||
Db      287  ACGTTTCGCAAATCGTTCCGCTAGACCATTGATCCAGAA---CAACTTGAGCAGGGTTCTATCA  343

Qy      467  TTTTGATTGGTTCTATGTCTGGTGGTGCCATTGTCAACGAACCTCAAAACCAAGTTGTCTCACA  526
             |||||| ||| |||  |   |||||||||||||| || |  |||  | |||| | ||   | |
Db      344  TCTTGATTGGCTCAAATGTTCAACTTGTCAAGAACAAATTGTCAAGGACGCCCAAAGTTATGTACA  403

Qy      527  ACAAGTCCAAGGCTGGGTGGTTATCGTTGGTGGTCCTTGGGAAATGGTTCTAAGT  586
             |||||||| || |||| ||||  | | |||   || |||| || || |||  |||
Db      404  ACATGTCAGAAGGTTGGATCCACTTGGTCAGATGCGTTTGGGCTGCGAAATGGGCCAAGT  463

Qy      587  ACAACATCAGAGTTAATTCTTTAAACCCAGGTTACATCTACGGTCCTTGACCAAGAATG  646
             |||| || |||| |||   |  |  |    ||| | ||| |||  |||  |  ||||
Db      464  ACAACATCAGAGTGATCAACACCTTATCACCAGGCTATATTTGACTCCTTGTTAACCAGAAACG  523

Qy      647  TTATCAATGGTAACGAAGAAATTGTACACAGATGGATCTTCTGGTATCCCACAACAACAACAAGA  706
             |||||| ||| ||| || ||||| || | | ||||  ||  ||  || | | || | |||
Db      524  TGATTTCTGGCCACACAGAGATGAAAGGAAGCCTCGGGAAATCCCAAGATCCCATGAAAGAGAA  583

Qy      707  TGTCCGACCAAAAGGAAATACATTCGGTGCTGCTGTTTTGTTACTTGCTTTCTGAATCTGCTGCTT  766
             ||||| ||||| | ||  |  |   | |   ||  |||| ||  ||| |||||| |||
Db      584  TGGCCGAACCCAAGGGAATTCGTGGGGTGGGTCCATCTTTATACTTGCAAGCGAGACTGCTTTCTT  643

Qy      767  CATACACTACTGTGTGCCAGCTTACTGGTGTGTGGTT  805
             |||||  || || |  |||  |||   ||||| |  |
Db      644  CCTACACTACGGGGCCCACAAATTTGGTTGTGTGGACGGAGGAT  682
```

```
RESULT 9
CR382121
CR382121_09/c
WPCOMMENT

Sequence split into 11 fragments  LOCUS CR382121 Accession CR382121
Fragment Name        Begin          End
CR382121_00              1         110000
CR382121_01         100001         210000
CR382121_02         200001         310000
CR382121_03         300001         410000
CR382121_04         400001         510000
CR382121_05         500001         610000
CR382121_06         600001         710000
CR382121_07         700001         810000
CR382121_08         800001         910000
CR382121_09         900001        1010000
CR382121_10        1000001        1062590
Continuation (10 of 11) of CR382121 from base 900001 (CR382121 Kluyveromyces lactis stra

Query Match    19.1%;  Score 155.8;  DB 8;  Length 110000;
Best Local Similarity  53.4%;  Pred. No. 1.1e-28;
Matches 473;  Conservative 0;  Mismatches 337;  Indels 75;  Gaps 4;
```

```
Qy        1  ATGACTGACTACATTCCAACTTTAGATTCGATGGCCACTTAACCATTGTCACAGGTGCC  60
             |||          ||| |||||      ||||      ||   |||| || | | |  |
Db    84309  ATTACTGATAGTTCTGCCTTCTTTTAGATTGGATGGACACGTAACCATCCTCACTGGTGCT  84250

Qy       61  TGTGGTGGTTTACTGAAGCTTTAATCAAGGGTTTGTTGGTGGACACGTAACCATCCTCACTTGT  120
             |||||||| | | |||  ||   || |||||||||  ||| || |   || |  | | || ||
Db    84249  GCCGGTGGGGTTGGCTCCACAACTTGCTCACACTTTGTCACAGGCAGTTGGTGTCTGACATTGCA  84190

Qy      121  TTGCTTGATATCGACCAAGAAAAAGAGACTGCTGCCAAAACAAGCCGAATACCACAAATACGCT  180
             ||| || ||||| ||| |   | | |  || |||  || | | || ||| | |||  ||||||
Db    84189  TTGATGGATCTTCTTCTTGCCTAGTTAAAACTGAAGAGAATGTTAAAACTGAAGAACTCTTC--  84133
```

```
Qy    181  ACTGAAAGAATTGAAGTTGAAAGAAGAGTTCCAAAGATGGGTTCATATGCCTGTGATATTTCT  240
                ||||                    ||| |       |||||||    |||||||  |||
Db  84132  ---CAGAAACAGAACAATATACCTTCCAGTTAAGATTTCCCTATGGGCTTGTGATAATATCG  84076

Qy    241  GATTCTGATACCGTTCACAAAGGTGTTTGCTCAAGTTGC----TAAGGATTTTGGTAAGTTG  297
                ||| |           ||| || |||||| |||||            ||| ||  |||||
Db  84075  AACCGAGATGAAGGTAGAAAATGGTGGTTAAGTCGAATTCCAAGTTCTCACAAGGGAATTGTA  84016

Qy    298  CCAATTGCACTTGGTTAACACAGCTGGTTACTGTGAAAAACTTCCCATGTGAAGAATTACCCA  357
                ||||    ||| |         |||    |||     |||||     |||  ||||  ||
Db  84015  CCACAAAAATTGGTACATACCGCTGGATTTTGTCAAAAACATCATGAGTACGAT  83956

Qy    358  GCCAAGAACGCTGAGAGAGATGGAAGGTTAACTTGTTGGGTTCCTTTTGTATGTTTTCTCAA  417
                ||||  |||||||   || |||| || |||    || |||||| |||  || |||||||||
Db  83955  GCTGAAAAGGCTCAAAAAATTGGTTAAGTTAACTTCTTTATACATTGCCAA  83896

Qy    418  GCCTTTGCTAAGCCATTGATCAAAAGAAGTA--------  448
                |||||| ||  | |||                          
Db  83895  TCTATGGCTAGACAACTATTGGGCTAGGAGAACTAGAAGAAAGTCTATCTCTGAGCACATT  83836

Qy    449  -----------------------TCAAGGGTGCTTCTCGTTGTTGTTTTTGATTGGTTCT  480
                                         |||  ||| | ||| ||| ||| |||| |||
Db  83835  GATGAATCTCAGGTGAAGCTATGCCTCAGTTCCCTGAAGCTTCATTCGTTCTCTAATCGGATCC  83776

Qy    481  ATGTCCGGTTTAATCGTTGTTAACACTGTTGTCTACAACATGTTGCTACCCTCAACATGTTC  540
                ||| ||| ||  |||  |||| || ||| |||||||  | ||||||  || ||| |||| |
Db  83775  ATGTCTGGTTGGTGCCATTGTCAACGATCCTCAAAACCAAGTTGTCTACAACCAAGTTCCAAGGCT  83716

Qy    541  GGTTCATTCACCTTAGTGAAATCACTAGCAGCAGAATGGGCTAAAATTGGTGTATTAGAGTT  600
                ||| ||||| | | |||||| |||  ||    |||||||||  |  | || ||| ||||||
Db  83715  GGTGTCATTCACCTTAGTGAAATCACTAGCAGCAGAATGGGCTAAAATTTGGTATTAGAGTT  83656

Qy    601  AATTCTTTAAACCCAGGTTACATCTTCCTTTGACCAAGAATGTTATCAATGGTAAC  660
                ||| || |   |||||| ||| | ||| | ||| |||  |||  |||| ||
Db  83655  AACGCCATTTCTCCAGGTTACATTGCTTGACCAAGAAGTCATCTCAATCG  83596

Qy    661  GAAGAA-----------TTGTACAACAGATGGAATCCTCGGTATCCCACAACAAAGAATGTCC  711
                ||||||              |||||||  |  ||| |   |||  |||||||||||
Db  83595  CAAGAAGGTGCTGCTTGCTCTACAAAAGAATGGACTAACAGAATCCCACTTGGTCGCCATGGCT  83536

Qy    712  GAACCAAAGGAATACATTGGTCGTTTGTTTTCGTTTCTGAATCTCCTGCTTCATAC  771
                || ||||| ||| ||| |||  |||| ||| |||||   |||| | |||| ||
Db  83535  GAACCAAAGGAATTTATCGGTCTTACCTTACTATCGAGATCTCGCTTCCAGTTAC  83476

Qy    772  ACTACTCTGGTGCCAGCTTACTGGTTGATGGTTGGGTTTCACTTCTTGG  816
                ||||||| |||   | ||| ||  |||  |||| ||| |||||||| ||
Db  83475  ACTACTCTGGTTGAAAATATGGTTGTTGGGATGGTGGTTACGTCGTGTTGG  83431
```

Fragment Name	Begin	End
CR382132_00	1	110000
CR382132_01	100001	210000
CR382132_02	200001	310000
CR382132_03	300001	410000
CR382132_04	400001	510000
CR382132_05	500001	610000
CR382132_06	600001	710000
CR382132_07	700001	810000
CR382132_08	800001	910000
CR382132_09	900001	1010000
CR382132_10	1000001	1110000
CR382132_11	1100001	1210000
CR382132_12	1200001	1310000
CR382132_13	1300001	1410000
CR382132_14	1400001	1510000
CR382132_15	1500001	1610000
CR382132_16	1600001	1710000
CR382132_17	1700001	1810000

Fragment Name	Begin	End
CR382132_18	1800001	1910000
CR382132_19	1900001	2010000
CR382132_20	2000001	2110000
CR382132_21	2100001	2210000
CR382132_22	2200001	2310000
CR382132_23	2300001	2410000
CR382132_24	2400001	2510000
CR382132_25	2500001	2610000
CR382132_26	2600001	2710000
CR382132_27	2700001	2810000
CR382132_28	2800001	2910000
CR382132_29	2900001	3010000
CR382132_30	3000001	3110000
CR382132_31	3100001	3210000
CR382132_32	3200001	3310000
CR382132_33	3300001	3410000
CR382132_34	3400001	3510000
CR382132_35	3500001	3610000
CR382132_36	3600001	3710000
CR382132_37	3700001	3810000
CR382132_38	3800001	3910000
CR382132_39	3900001	4003362

Continuation (4 of 40) of CR382132 from base 300001 (CR382132 Yarrowia lipolytica chromo

Query Match   16.0%; Score 130.8; DB 8; Length 110000;
Best Local Similarity 56.3%; Pred. No. 2.3e-22;
Matches 287; Conservative 0; Mismatches 217; Indels 6; Gaps 2;

```
Qy    307  TTGGTTAACACAGCTGGTTACTGTGAAAAACTTCCCATGTGAAGAATTACCCAGCCAAGAAC  366
                |||||||||||||||||||||||||  |||||||||||| || ||||||||||||||||
Db  44725  TTGGTCAACTGTGCGCCGGATTCTGCGGAGAATAATGACTGGCCTTTGATTTATCCCAACCCCAG  44666

Qy    367  GCTGAGAAGGATGGTGAAGGTTAACTTGTTGGGTTCTTTTGTATGTTTTCTCAAGCTTTGCT  426
                |||   |||| |||| ||||||||  || |||  |  ||| || || ||||||| || ||
Db  44665  GTCAAGCGACTGCTGGACGTCAAACCTCATGGGATCCTACACAAACTTCGCCTACACGAGGTGGCC  44606

Qy    427  AAGCCATTGATCAAAGAAGGTATACAAAGGGTGCTTCGTTCGTTGTTTTTGATTGGTTCTCT  486
                |||||||| | |||||     ||  |||||  | |||||  | |||||||| |||||||| |
Db  44605  AAGTCGCCTTGTCCTGACG---AGTCCTCGGATCTCTGATTCTGATTCCATGAGT  44549

Qy    487  GGTGCCATTGTCAACGATCCTCAAAACCAAGTTGTCTACAACATGTTGCTACCCTCGGTGTT  546
                |||||||| |||  | || | ||||||| ||| || |||| ||||| ||| ||| | | |||
Db  44548  GGCTCCATTGTCAACGGACCCCCAGACCCCAGCAGGGTAGTTCCAAGGCAGGTGTC  44489

Qy    547  ATCCATTTGGCTAACGACTTTGGGCTTGTGAATGGGCGCTAAGTACAACATCAGAGTAATTCT  606
                ||||||||||  |||| |||| ||| | | ||| | ||| |||| ||| | |||| || |||
Db  44488  ATCCACATGGCCACAAGTCTCTCGTGGGCCCGAGTGGGCCCCAGTACCAACACTCAGACT  44429

Qy    607  TTAAACCCAGGTTACATCTACGGTGTCCTTTGACCAAGAAGTGTTATCAATGGTAACGGAAGAA  666
                ||||||||||| |||| |||    |||||| |||||| ||||  | ||||| ||| ||| |||
Db  44428  CTGTCTCCCGGCTACATTCTTACTCCCTCGTACCCTCGACCCCGTCACATCATCGAGACTGACGGAGG  44369

Qy    667  TTGTACAACAGATGGAATCCTCGGTATCCCACAACAAAGAATAGTTCCGAACCAAAGGAATAC  726
                |||||||||||| |||| | |  || |||||||||||||  | ||   |||||||||||| |
Db  44368  CTCCGAAACGACTGGGGAGCGGAGAATTCCTTTCCGAACGAATGGCCGGAACCAAAGGAGAGTTT  44309

Qy    727  ATTGGTTGCTCGTTTTTGTTACTTGCTTTCGTTGAATCTGCTGCTTCATACACTACTACTGG  786
                |||||| |||  || |||  ||||||| || ||| | | | |  || |||| || || |||
Db  44308  GGAGGCGCCCTTATTGTCTTCATGGCCTTCCGA---CGCCTCCAGCGCTACATGGACGTGGTGCCAGC  44252

Qy    787  TTACTGGTTGAATGGTGGTTGGGTTTCACTTCTTGG  816
                |||| ||||| ||| ||| |||| ||||| |||| ||
Db  44251  CTCAATTGTCATGGAGGGTTACACCAATCTGG  44222
```

Fragment Name	Begin	End
CR382133_00	1	110000
CR382133_01	100001	210000
CR382133_02	200001	310000

CR382133_03   300001   410000
CR382133_04   400001   510000
CR382133_05   500001   610000
CR382133_06   600001   710000
CR382133_07   700001   810000
CR382133_08   800001   910000
CR382133_09   900001   1010000
CR382133_10   1000001   1110000
CR382133_11   1100001   1210000
CR382133_12   1200001   1249565

Continuation (12 of 13) of CR382133 from base 1100001 (CR382133 Debaryomyces hansenii ch

Query Match   11.3%; Score 92.6; DB 8; Length 110000;
Best Local Similarity 52.5%; Pred. No. 1.1e-12;
Matches 309; Conservative 0; Mismatches 259; Indels 21; Gaps 4;

```
Qy    221 CAATGCCTGTGAATATTTCTGATTCGATACCGTTCACAAGGTGTTTGCTCAAGTTGCTA 280
          ||| |||     ||  ||  |||||   |||| |||  || |||| |||||||||| ||
Db  99071 CTTATAAATGTCGTTGTTGATTAATTTTGAAGAAGTAGAAAAACTGTAAATCAAATAGTTGA 99012

Qy    281 AGGATTTTGGGTAAGTTGCCATTGCACTTGGGTTAACACACTGGTTACTGTGAAAAACTTCC 340
          ||||||| ||   |  ||| ||   |  ||| |||| |   |  |  |||||||   ||| ||
Db  99011 AGGATTTTTGGAACAAATTGATATATTCGTTGCTAATGCTGGGTTTGGCTTGGAAATCCTGCC 98952

Qy    341 CATGTGAAGATTACCCACGCCAAGAAACGCTGAGAAGAAGCGTTAACTGTTTGGGTT 400
          ||||| |||| | |||| | |||||| ||||||||| |||  |||||| |  || ||
Db  98951 CAGTAAGCGCAGTCTTCATTAGATGCCTACCACAATGTGATTGAAAAGTTAACTTAGACGGGCG 98892

Qy    401 CTTTGTATGTTTCTCAAGCCCTTTGCTAAGCCATTGATTCAAAAGAAAATCCTGCTT 460
          |||||| |||  || ||||  | |   | || |  | | |||| || |||||  |||
Db  98891 TTTTACTATGC------AGCCAAATGTATAGGACCAATTTTCGGAAAGAATGGGAAAAGGAT 98838

Qy    461 CTGTGTTTTGATTGGTTCTATGTCTCGGTCGTCAAAAACCAAGTTG 520
          |||||||| ||| | | |  ||||||| ||| |||| ||||| |||
Db  98837 CAATTATAATTACTGCTTCCCAAGCTGACATATAGTCAACATTAAGTCAAGCTG 98778

Qy    521 TCTACAACATGTCCAAGGCTGGGTGTTATCCATTGGCTTGTGAATGGG 580
          ||||||| ||||||||| ||||| |||||||||||||| || ||| |||
Db  98777 CCTTATAATGCAAGTAAAGCAGCTCTTATCCAATTGGGAAAAATCACTAGCAATTGAGTCGGGG 98718

Qy    581 CTAAGTACAACATCAGAGTTAATTCTTTAAACCCAGGTTACATCTACAGGTTCCTTTGACCA 640
          |||||||| | |||| |||| |   |  ||    || ||||||  | || || |  |||| ||
Db  98717 TTAGAT--TTGCAAGAGTTAATACTATATCACCAGGTTATAT------TTTGACCG 98670

Qy    641 AGAATGTTATCAATGGTAACGAAGAAAATGTACAACAGATGGATCTCGGTATCCCACAAC 700
          |||| ||||| ||  |  ||||||||||||| | ||| ||||||| |||| |||||||||
Db  98669 AGATTGCGAAACATGTTCCCTTTGATGTGAAGTCAAAAATGGGCCTTTAATTCCAATGG 98610

Qy    701 AAAGAATGTCCGAACCAAAAGGAATAACATTGGTTGGTGCTGTTTTGTTGTACTGGTGCCAA 760
          ||| |||  ||||| |||||  ||| | |||| |  |   | |||| |||||  |||| |||
Db  98609 GGAGAGAGGGCTCGCCACAAGAATTAGTCGGAGCATTTTGTATTTTTTGTAATCTG 98553

Qy    761 CTGCTTCATACACTACTGGTGCCAGCTTACTGGTTGTTGTTTCAC 809
          ||||||||| || |||| |||| | ||||  |||   |   |||||
Db  98552 CATCTACATTCACTACTGGGTGCAGACTTAATAATTCGATGGCCGGATATAC 98504
```

Continuation (7 of 21) of CR382137 from base 600001 (CR382137 Debaryomyces hansenii chro

Query Match   11.2%; Score 91.4; DB 8; Length 110000;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;
Matches 331; Conservative 0; Mismatches 276; Indels 24; Gaps 5;

```
Qy    182 CTGAAGAATTGAAGTTGAAAAGAAGTTCCAAAGAAGATGGGGTTCATATGCCGTGATAATTTCTG 241
          |||||||||| || ||||| || ||  || || |||||| ||||| |   |||  |||| |||
Db  92817 CTGAATACTTAACCAAAGAACCTACCGGAGTGAAAATCCAAGCATACAAAGTGTAAATATTTCGG 92876

Qy    242 ATTCTGATACCGTTCACAAGGTGTTTGCTCAAGTTGGCTAAGGATTTGGTAAGTTGCCAT 301
          ||| ||||   ||| ||||||  | || || |  ||  ||| ||||| ||| |||| |||
Db  92877 ACCCAGAAGAGATGTTGGAAAAAGGTGCGCTAAGGGTTGCCAAGGAGGAAATTTGGAGACG 92936

Qy    302 TGCACTTGGTTAACACAGCTGGTTGTTACTGTGAAAACTTCCCATG---TGAAGATTACCCAG 358
          |||| || | |||  |  ||| || | ||||| |||| ||||| ||   ||||| |||| |||
Db  92937 TTTTTGTCGGTCAATGTCGTGTGCCATGGAACAGAAATTTGAACTTTCCGGCACTATTGACG 92996

Qy    359 CCAAGAACCCTGAGAAGATGGGTGAAAGGGTTAACTTGTTCGGGTTCTTGTATGTGTTTCAAG 418
          |||| || | | ||| | |||||||  |||||| ||| | |||| ||   | | |  || |||
Db  92997 ACGATGCCTGGAAAAAAATTGTAGATTTGGACTTGGAGTTGTTACTACTGAGA 93056

Qy    419 CCTTTGCTAAGCCATTGATCAAAGAGAAGGTGTATCAAGGGCCTTCGTTGTTTTGATTGGTT 478
          ||||||||||| |||| || | ||| | |||| | ||||  ||  | |||| | ||| ||||
Db  93057 CTGTCGGCAAGATTTTCAAGGAGCACACGGAA------AGGGTTCTTCTTGGTTTAACAGCAT 93110

Qy    479 CTATGTCTCGGTGCCATTGTCAAGGATCCTCAAAACCAAGTTGTCTACAACATGTCCAAGG 538
          ||||||| | || |||||| |  | | | |||||| |||||||| |||| |||||||||||
Db  93111 CAATGTCTGGCCACATTGTTGAATGTTCCACAAAATGCAAGGCTTCCATACAACAGCGCTAAGG 93170

Qy    539 CTGGTGTTTATCCATTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGGTACAACATCAGAG 598
          |||| ||| |||| ||||| |||||||||||| ||| || || ||| || |||| | ||||
Db  93171 CGGGTGTATTGCACTTGGGTAAGTCCTTTAGCATACGAAGTGGGGTTCCATTGCT---AGAG 93227

Qy    599 TTAATTCTTTAAACCCAGGTTACATCTACAGGTCCTTTGACCAAGAAGAATGTTATCAATGGTA 658
          ||||||| |  ||||||||||||||| ||| ||   || ||||||| | || | | |||||
Db  93228 TTAACACCATTTCCAGTTACAT------TGCTACAGAGAATTTTCGGAGTTTGTTC 93278

Qy    659 ACGAAGAATTGTACAACAGATGGATCTCTGGGTATCCCACAACAAAGAATGTCCGAACCAA 718
          ||||| |||||| |||||||| | ||| |||||||||| |||| ||||| || |||| |||
Db  93279 CTGCCGAAGTCAAGGCTCAAAATGGGTGGCAATTGAATCCCATTAGGAGAGACACTTGCTC 93338

Qy    719 AGGAATACATTGGTGGCTGCTGTTTTGTTGTACTTGCTTCGTTCATCACTACTACTG 778
          |||||||||||| |   || || | || ||| |  |  |  || |||||||| |||| |
Db  93339 AAGAGTTGGTTGGTGCTGCTTACTTGATTATTTTGCATCCGA---TGCTTCTACTTACACTG 93395

Qy    779 GTGCCGCACTTACTGGTTGATGGTGGTTTCAC 809
          ||||||| || | ||||| |||||||| ||||
Db  93396 GCTCGGACTTATTGGTTGATGGTGGTTTACTC 93426
```

RESULT 13

AF002134

LOCUS   AF002134   4293 bp   DNA   linear   PLN 26-JUN-1998
DEFINITION Candida albicans Sou2p (SOU2), Sou1p (SOU1) and Vma8p (VMA8) genes,
            complete cds.
ACCESSION   AF002134
VERSION   AF002134.1 GI:2183241
KEYWORDS   .
SOURCE   Candida albicans
ORGANISM   Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.

RESULT 12
CR382137_06
WPCOMMENT
Sequence split into 21 fragments   LOCUS CR382137 Accession CR382137

Fragment Name   Begin   End
CR382137_00   1   110000
CR382137_01   100001   210000
CR382137_02   200001   310000
CR382137_03   300001   410000
CR382137_04   400001   510000
CR382137_05   500001   610000
CR382137_06   600001   710000
CR382137_07   700001   810000
CR382137_08   800001   910000
CR382137_09   900001   1010000
CR382137_10   1000001   1110000
CR382137_11   1100001   1210000
CR382137_12   1200001   1310000
CR382137_13   1300001   1410000
CR382137_14   1400001   1510000
CR382137_15   1500001   1610000
CR382137_16   1600001   1710000
CR382137_17   1700001   1810000
CR382137_18   1800001   1910000
CR382137_19   1900001   2010000
CR382137_20   2000001   2037969

```
REFERENCE   1 (bases 1 to 4293)
  AUTHORS   Janbon,G., Sherman,F. and Rustchenko,E.
  TITLE     Monosomy of a specific chromosome determines L-sorbose utilization:
            a novel regulatory mechanism in Candida albicans
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (9), 5150-5155 (1998)
   MEDLINE   98226783
   PUBMED    9560244
REFERENCE   2 (bases 1 to 4293)
  AUTHORS   Janbon,G., Rustchenko,E. and Sherman,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAY-1997) Department of Biochemistry and Biophysics,
            University of Rochester, 601 Elmwood avenue P.O. Box 712,
            Rochester, NY 14642, USA
FEATURES             Location/Qualifiers
     source          1..4293
                     /organism="Candida albicans"
                     /mol_type="genomic DNA"
                     /strain="SOl7"
                     /db_xref="taxon:5476"
                     /chromosome="4"
     gene            356..1199
                     /gene="SOU2"
     CDS             356..1199
                     /gene="SOU2"
                     /note="short-chain alcohol dehydrogenase homolog"
                     /codon_start=1
                     /transl_table=12
                     /product="Sou2p"
                     /protein_id="AAC24462.1"
                     /db_xref="GI:2183242"
                     /translation="MSKETTSYTNAKLGPLPTKAATIPDNILDAFSLKGKVASVTGSS
                     GGIGWAVAEGYAQAGADVAIWNSHPADDKAEYLTKYGVRSKAYKCNVTDFQDVERV
                     VKQIESDGTIDIFVANAGVAWTEGKPFDVKGVVDKMRKVVDVDLNSYYYCAHVGPIF
                     RKKGKGSPIFTASMSASIVNVPQLQAAYNARKAGVKHLSKSFTVEWAPFARVNSVSPG
                     YIATHLSEFADPDVKSKWLQLTPLGREAKPRELVGAVLYLASDAASYTTGADLAVDGG
                     YTVI"
     gene            complement(2063..2909)
                     /gene="SOU1"
     CDS             complement(2063..2908)
                     /gene="SOU1"
                     /note="short-chain alcohol dehydrogenase homolog"
                     /codon_start=1
                     /transl_table=12
                     /product="Sou1p"
                     /protein_id="AAC24463.1"
                     /db_xref="GI:2183243"
                     /translation="MSEEIISFTNPALGPLPTKAPQLPSNVLDLFSLKGKVASVTGSS
                     GGIGWAVAEAFAQAGADVAIWNSKPADAKAEYLTEKYGVRAKAYKCNVTDNDVSKV
                     INEIEKDFGTIDIFVANAGVAWTDGPEIDVQGYDQWKKIVDCDLNGYYYCAHTVGQIF
                     KKNKSGSLIITSMSGIVVNIPQLQAYYNAAKAACTHLAKSLSVEWASPGARVNSISP
                     GYILTDIADFADPEMKKKWWLQLTPLGREGLPQELVGAYLYLASNASTYTTGSNIAVDG
                     GYTCP"
     gene            complement(3073..3877)
                     /gene="VMA8"
     CDS             complement(3074..3877)
                     /gene="VMA8"
                     /note="vacuolar H+-ATPase subunit B homolog"
                     /codon_start=12
                     /transl_table=12
                     /product="Vma8p"
                     /protein_id="AAC24464.1"
                     /db_xref="GI:2183244"
                     /translation="MSGAGNREQVFPRMTLGVMKSKLKGAQQGHSLLKRKSEALTKR
                     FRDITQRIDDAXRKMGRNVMQTAAFSLAEVQAYATGDNISYQVQGESVQKARFTVKAKQEN
                     VSGVFLFTFDSHINEDVNDFKLTALARGGDQVQXAKLIYSKAVELVELASLQTAFII
                     LDEVIKITNRRVNAIEHVIIPRTENTIAYINGELDEMDREEFYRLKKVQEKKQEAAAA
                     AEQEEALAKAKAEGADTDELAIQQDVEALDIKADKEEVDILQEKEDDVIF"

ORIGIN
```

```
Query Match     10.7%;  Score 87.2;  DB 8;  Length 4293;
Best Local Similarity    49.7%;  Pred. No. 2.9e-11;
Matches  314;  Conservative  0;  Mismatches 303;  Indels 15;  Gaps   3;
```

```
Qy    182 CTGAAGAATTGAAGTTGAAGAAGTTCCAAAAGATGGGTTGTCATATGCCTTGATATTTCTG 241
          |||| ||| || || |||| || |||  ||  |||| ||||| |  || || | |||| ||
Db   2286 CTGAATATTTAACTGAAAAAATATGGGTGTCAAAGCCAAAGCTTATAAAATGTAAGTAACTG 2345

Qy    242 ATTCTGATACCGTTCACAAGGTGTTGTCTCAAGGTGTTGCTCAAGTTGCTAAGGATTTTG 301
          | | |||| |  ||| ||||| | || || ||||| || || ||| || | ||||||  |
Db   2346 ATCCTAATGATGTTTCTAAAGTGTAACTGTGATTAATGAAATGAAAAGAGAATTCGGGTAC 2405

Qy    302 TGCACTTGGTTAACACAGCTGGTTACTGTGAAACTTCCCATGTGAAGATTACCCAGCA 361
          || | ||| ||||||| |||| || || || ||||| |||| ||  ||||| ||  ||
Db   2406 TATTTTGGTTGCTAAATCCTGGAGTTGCAATGGGACTGATGGACCAGAAATTGATGTTCAAGGCT 2465

Qy    362 AGAACGGCTGAAGAAGATGGGTGAAGGTTAACTTGTTGGGTTCTTTGTAGTTGTTCTCAAGGCCT 421
          || ||| | ||||| |  ||| ||| ||||||   |||| ||  |  ||| || |  | |||
Db   2406 ATGATCAAATGAAAAAAGATCGTTGATTGTGATTTAAATGAGAGTTTATTAATTGT---GCTC 2522

Qy    422 TTGCTAAGCCATTGATCAAAGAGGTATCAAGGGTCGCCTTCGTGTTGTTCTTGGTTCTA 481
          || ||||| |  ||||| |||| ||| |  ||||| || || ||||| ||  | |||| |
Db   2523 ATACCGTGGGACAAAATCTTTAAAGAGAATAAAATCTGGTTCATTAATTATTAATTCAAGAA 2582

Qy    482 TGTCTGGTGCCATTGTTCACGGATCCTCAAAACCAAGTTGTCTAACACATGTCCAAGGCTG 541
          |||| ||| |||| ||||||| |||| ||||| ||||||  ||||||||| |  | ||||||
Db   2583 TGTCAGGGACAAATTGTTAAAATCCCTCAAATTACACAGCTCCTTATAAATGCTGCTAAAAGCTG 2642

Qy    542 GTGTTATCCATTTGCTAAGACTTTTGGCCTTGTGAATCGCTGCTAAGTACACATCAGAGTTA 601
          | |||| || | || ||| ||||| ||||| |||||  ||| ||| |||| || |  | |||
Db   2643 CATGTACTCATTTAGCCAAATCATTGAGTGTGGAATGGGCCGCAGTTTTGGGTGCTAGAGTAA 2702

Qy    602 ATTCTTTAAACCCAGGTTACATCTACGGGTCCTTTGACCAAGAAATGTTATCAAATGGTAACG 661
          ||| ||| ||||||||  ||| || |||  ||||||| || ||| |||  | |  || ||| |
Db   2703 ATTCAATTTCTCCAGGGTATATA---------TTGACTGATAATTGCTGATAATTGCTGATC 2753

Qy    662 AAGAATTGTACAACAGATGGATCTCTGGTATCCCACACAACAAAGAAATGTCCGAACCAAAGG 721
          ||||| ||| |||||| ||||  ||| || |||||| |||| |||||||||| |  |||||||
Db   2754 CAGAAATGAAAAAAGAGAAATGTGGCAAATTGACACCCTTTGGGAGAGAAGGATTACCAACAAG 2813

Qy    722 AATACATTGGTGCTGCTTGTTTGTTACTTGCTTCGTCATCACTACTGGGTG 781
          || | | ||| || || ||| || ||||||| |||| ||||||||||  ||
Db   2814 AATTAGTGGGGGCATATTTATACTTGGCCTC---AAATGCATCAACTTAATACTACTGGGTT 2870

Qy    782 CCAGTTACTGGGTTGGATGGGGGGTTCACTTCT 813
          |||||||||||| | ||||||||| ||||| ||
Db   2871 CAAATATTGCTGTTGATGGGGGTTATACATGT 2902
```

```
RESULT 14
AC105434/c

LOCUS       AC105434               126105 bp    DNA     linear   HTG 20-MAR-2004
DEFINITION  Magnaporthe grisea chromosome 7 clone 18L14, *** SEQUENCING IN
            PROGRESS ***, 2 ordered pieces.
ACCESSION   AC105434
VERSION     AC105434.2  GI:45597494
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Magnaporthe grisea (anamorph: Pyricularia grisea)
  ORGANISM  Magnaporthe grisea
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
REFERENCE   1 (bases 1 to 126105)
  AUTHORS   Thon,M.R., Mitchell,T., Brown,D.E., Diener,S., Taro,A., Pan,H. and
            Dean,R.A.
  TITLE     The sequence of Magnaporthe grisea chromosome 7
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 126105)
  AUTHORS   Dean,R.A. Dr, Mitchell,T. Dr, Thon,M. Dr and Brown,D.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JAN-2002) Plant Pathology - Fungal Genomics
            Laboratory, North carolina State University, 840 Main Campus Drive,
            Raleigh, NC 27606, USA
REFERENCE   3 (bases 1 to 126105)
  AUTHORS   Thon,M.R., Mitchell,T., Brown,D.E., Diener,S., Taro,A., Pan,H. and
            Dean,R.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2004) Center for Integrated Fungal Research,
```

```
COMMENT     North Carolina State University, Plant Pathology - 840 Main Campus
            Dr, Raleigh, NC 27606, USA
            On Mar 20, 2004 this sequence version replaced gi:18071329.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 2 contigs. Gaps between the contigs
          * are represented as runs of N. The order of the pieces
          * is believed to be correct as given, however the sizes
          * of the gaps between them are based on estimates that have
          * provided by the submitter.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          *     1       86195: contig of 86195 bp in length
          *     86196   86395: gap of unknown length
          *     86396  126105: contig of 39710 bp in length.

FEATURES             Location/Qualifiers
     source          1..126105
                     /organism="Magnaporthe grisea"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:148305"
                     /chromosome="7"
                     /clone="18Li4"

ORIGIN

Query Match        10.6%;   Score 86.8;  DB 2;   Length 126105;
Best Local Similarity   53.6%;   Pred. No. 3.3e-11;
Matches  181;  Conservative   0;  Mismatches 157;  Indels  0;  Gaps  0;

Qy     464 TTGTTTTGATTGGTTCTAAGTCTCGGTGCCATTGGTCAACCGATCCTTCAAAAACCAAGTTGTTGTCT 523
           |||||| ||||| || ||| |||||||| |||| || | || ||| | ||| ||| || |||| || |
Db   56680 TGGTCTTCATCGGCAGCATGAGCGCGGCGCCATCGTCGACCGTGCCGACGCCGCAGGCCCCCT 56621

Qy     524 ACAACACGTCCAAGGCTCGGTGTTATCCATTTGGCTAAGACTTTGGCTTGTGTGAATGGGCTA 583
           ||| || |||| ||| || | || | | ||||| | || |||| ||| || ||| || |||
Db   56620 ACAACAGCCGTCAAGGCCGCGTCGTGCCCACCTGGCCGCCCGTCGGACTGGGGGCGGC 56561

Qy     584 AGTACAACATCAGAGTTAATTCTTTAAACCCAGGTTACATCTACGGGTCCTTTTGACCAAGA 643
           |||||||| || ||| || ||| | | |||||||| ||||||| |||||| ||| |||||||
Db   56560 ACGCCGGCATCCGGGTCAACTGCATCTCTCCCGGCTCAATGCTCACCGGGTCGACCAAGA 56501

Qy     644 ATGTTATCAAATGGTAACGAAGAAATTGTACAACAGATGGATCTTCTGGGTATCCCCACACAACAAA 703
           | || | || || || | |||| ||| || || ||| |||| || || || |||| ||| || ||
Db   56500 AGATACTCGACGAGACAAGCCCGACCTCAAGAAGAAGTGGACCAGCCTCATCCCCAGGGCA 56441

Qy     704 GAATGTCCGAACCAAAAGAGAAATACATTGGTCGTGCTGCTGTTTTGTACTTGGCTGAATCTGCTG 763
           ||||| |||| ||| ||| | ||||| |||| |||||| || |||| ||| || || | || |||
Db   56440 GGATGGGGCCCCCAGAGAGACCTGATGGGGTCCCGTGACCTTCCGTGTGCGGACGCGGTCCG 56381

Qy     764 CTTCATACACTACTGGTGTGCCAGCTTTACTGGTTGATGGT 801
           ||||||||| || | || |||| |||| || |||||| |||
Db   56380 TCTACGTCACCGGCGCTGACATCCGTCGCGATGCGGGT 56343
```

```
Query Match        10.6%;   Score 86.6;  DB 6;   Length 927;
Best Local Similarity   52.0%;   Pred. No. 4.3e-11;
Matches  328;  Conservative   0;  Mismatches 279;  Indels  24;  Gaps  5;

Qy     182 CTGAAGAATTGAAGTTGAAAGAAGTTCCAAAGATAGGGGTTCATATAGCCGTGATATTCTG 241
           ||||| ||||| || ||| | || |||| | || ||| ||| || ||| || || | |||
Db     308 CTGAATAATTTAACCAAGACTTATCGGGGTTAAATCCAAAGCATACAAAGTAAATGTTACTG 367

Qy     242 ATTCTGATACCGTTCACAAGGCTGTTTGCTCAAGGCTGTAAGGATTTGGTAAGTTGGCCAT 301
           | || || || || | ||| ||||| || |||| || || |||| ||| ||| || ||
Db     368 ATTTCCAAGATGTTGAAAAAGTTGTCAAACAAATTGAACTGGGATTTCGGTACCATTGATA 427

Qy     302 TGCACTTGGTTAACACAGCTGGTTACTG---TGAAAACTTCCCATGTGAAGATTACCCAG 358
           || || ||| || ||| ||| || |||    ||||| ||| | || ||| | || ||||
Db     428 TCTTTGTTCCCAATGCTGCTGGTTGTTCGCCTGACCGAAGGGCCCGAAAATCGATGTCAAGGGAG 487

Qy     359 CCAAGAACGGCTGAGAAGAGATGGTGAAGGTTAACTTGTTGGGGTTCTTGTATGTTTCTCAAG 418
           ||||| |||||| || ||| || ||||| ||||| | ||| |||| || || ||| ||||
Db     488 TCGACAAATGGACAGAAGTTGTTGATGTTAAAGAGTGTTTAAATATTGTGTCCATCATG 547

Qy     419 CCTTTGCTAAGCCATTGATCAACAAGAAGGTATCAAGGGTGCTTCTGTGTTGTTTGATTGGTT 478
           |||| || ||| || ||| ||| ||| || || ||| || || || ||| || || ||||
Db     548 TTGTTGGTCCAATTTTCAGAAAAAAGGGTA------AAGGGTCAATTCATTTCATTTCACTGCCA 601

Qy     479 CTATGTCTCGGTGGCCATTGGTCAACGATCCTCAAAAACCAAGTTGTGTCTACAACATGTCCAAGG 538
           || || | ||| || ||| || | ||| | ||| || ||| || || |||| || ||| ||
Db     602 GTATGTCGGCGCTTCAAATTGTTAATGTCCCACAATTACAACGGCTGCTAAAG 661

Qy     539 CTGGTGTTATCCATTGGCTAAGACTTTGGCTTGTGTGAATGGGCTAAGGTAACACATCAGAG 598
           |||| ||| | ||| ||| ||| || ||| || ||| || || ||| || || ||| ||
Db     662 CTGGGGTCAAACATTTGTCAATCATTGGTGTCAAGATGTTGAAGATGGGGACCATTGCT---AGAG 718

Qy     599 TTAATTCTTTAAACCCAGGTTACATCTACGGGTCCTTTTGACCAAGAATGTTATCAAATGGTA 658
           || ||| | ||| ||||||| || || ||| || || ||| ||| || || |||| || |
Db     719 TCAATTCTTGTTCTCCAGGTTACATGCTACTCATCTTAGTGGAATTGGCTG--------- 769

Qy     659 ACGAAGAAATTGTACAACAGATGGATCTTCTGGGTATCCCACAACAAAGAATGTCCGAACCAA 718
           | |||| ||| || ||| ||| || ||| |||||| || || ||| ||| || ||| || ||
Db     770 ATCCGGATCGTCAAGAGTAAAAATGGTTGCAACTTACACACCACTTGGTAGAGAAGCCAAACCAA 829

Qy     719 AGGAATACATTGGTCGTGCTGCTGTTTTGTACTTGGCTGAATCTGCTGCTTCAGAAATCACTACTG 778
           || |||| || ||| || || || |||| ||| || || |||||| || || ||| |||| ||
Db     830 GAGAGCTTGTTGGTGCCTACTTGTTGTATTTGGGCTTCCGA---TGCTGCACTCTTATACAACTG 886

Qy     779 GTGCGCAGCTTTACTGGTTGATGGTTGGGTTTCAC 809
           ||| ||| ||| | ||| ||| || ||||| |||
Db     887 GTGCGTGATCTGCTTGTTGATGGTGGTTTGTACAC 917
```

Search completed: March 2, 2005, 15:57:30
Job time : 3861 secs

```
RESULT 15
LOCUS       AR547359                 927 bp    DNA       linear  PAT 08-OCT-2004
DEFINITION  Sequence 2490 from patent US 6747137.
ACCESSION   AR547359
VERSION     AR547359.1  GI:53940534
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 927)
  AUTHORS   Weinstock.K.G. and Bush.D.
  TITLE     Nucleic acid sequences relating to Candida albicans for diagnostics
            and therapeutics
  JOURNAL   Patent: US 6747137-A 2490 08-JUN-2004;
            North Carolina State University, Plant Pathology; 840 Main Campus
FEATURES             Location/Qualifiers
     source          1..927
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```